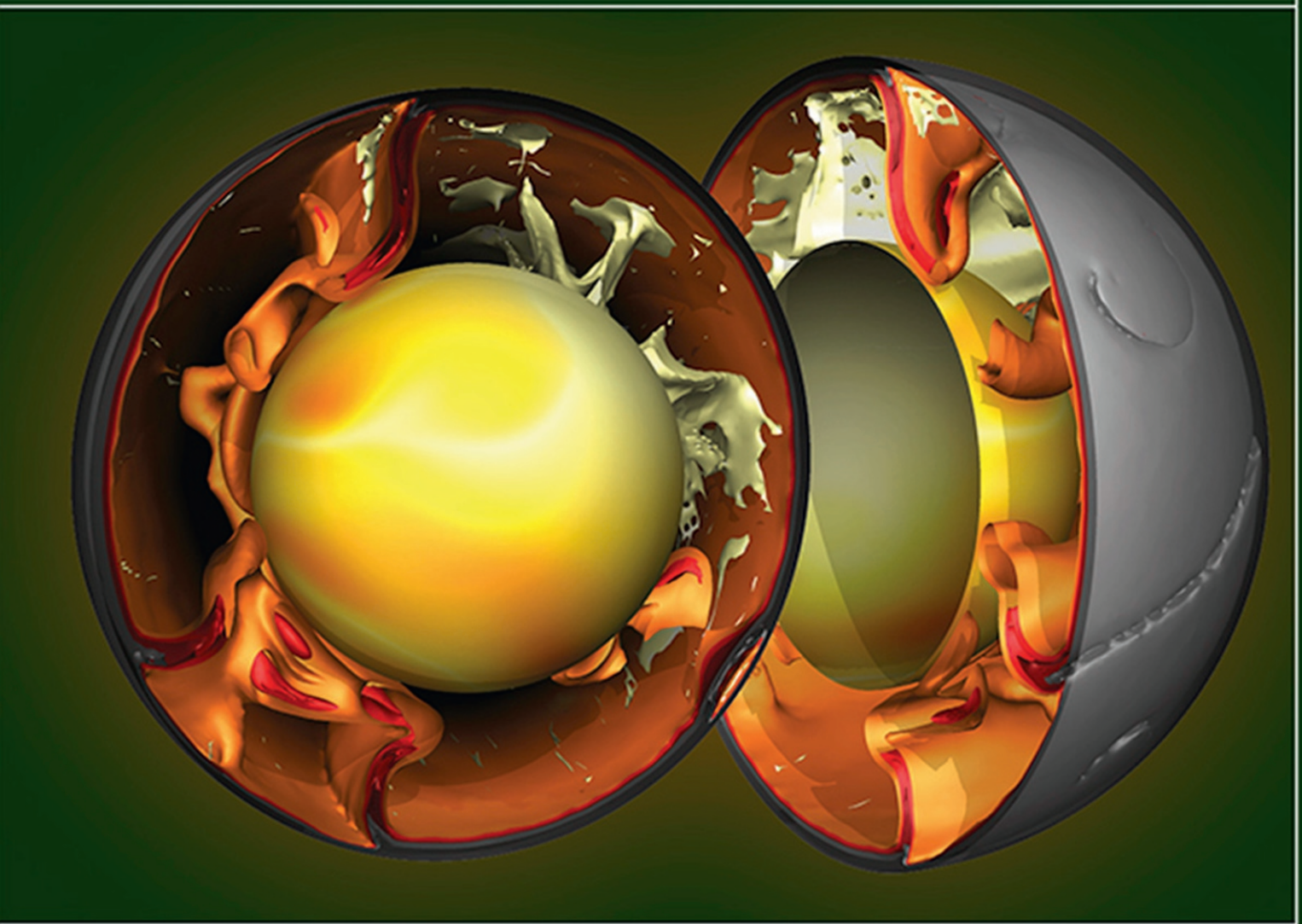
SECOND EDITION

TREATISE ON GEOPHYSICS

Volume 4

EARTHQUAKE SEISMOLOGY



VOLUME EDITOR: HIROO KANAMORI

EDITOR-IN-CHIEF: GERALD SCHUBERT

TREATISE ON GEOPHYSICS

SECOND EDITION

This page intentionally left blank

TREATISE ON GEOPHYSICS

SECOND EDITION

EDITOR-IN-CHIEF

Gerald Schubert

Distinguished Professor Emeritus, Department of Earth, Planetary and Space Sciences, University of California, Los Angeles

VOLUME 4

EARTHQUAKE SEISMOLOGY

VOLUME EDITOR

Professor Emeritus Hiroo Kanamori

Seismological Laboratory, California Institute of Technology

ELSEVIER

AMSTERDAM • BOSTON • HEIDELBERG • LONDON • NEW YORK • OXFORD
PARIS • SAN DIEGO • SAN FRANCISCO • SINGAPORE • SYDNEY • TOKYO

Elsevier
Radarweg 29, PO Box 211, 1000 AE Amsterdam, The Netherlands
The Boulevard, Langford Lane, Kidlington, Oxford OX5 1GB, UK
30 Corporate Drive, Suite 400, Burlington MA 01803, USA

Copyright © 2015 Elsevier B.V. unless otherwise stated. All rights reserved.

Cover image created for this work by Fabio Crameri, University College London.
The figure depicts the temperature field from a convection calculation using the code STAG-YY, written by Paul Tackley.

The following chapters are US Government works in the public domain and not subject to copyright:

1.11 Crust and Lithospheric Structure - Global Crustal Structure; 3.09 Earth Rotation Variations – Long Period; 4.07 The Role of Fault-Zone Drilling; 5.11 The Time-Averaged Field and Paleosecular Variation; 10.21 Exoplanetary Geophysics: An Emerging Discipline; 11.12 Tools and Techniques: Gravitational Method.

No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, or any information storage and retrieval system, without permission in writing from the publisher. Details on how to seek permission, further information about the Publisher's permissions policies and our arrangements with organizations such as the Copyright Clearance Center and the Copyright Licensing Agency, can be found at our website: www.elsevier.com/permissions.

This book and the individual contributions contained in it are protected under copyright by the Publisher (other than as may be noted herein).

Notices
Knowledge and best practice in this field are constantly changing. As new research and experience broaden our understanding, changes in research methods, professional practices, or medical treatment may become necessary.

Practitioners and researchers may always rely on their own experience and knowledge in evaluating and using any information, methods, compounds, or experiments described herein. In using such information or methods they should be mindful of their own safety and the safety of others, including parties for whom they have a professional responsibility.

To the fullest extent of the law, neither the Publisher nor the authors, contributors, or editors, assume any liability for any injury and/or damage to persons or property as a matter of products liability, negligence or otherwise, or from any use or operation of any methods, products, instructions, or ideas contained in the material herein.

Library of Congress Cataloging-in-Publication Data
A catalog record for this book is available from the Library of Congress

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library

ISBN (print): 978-0-444-53802-4

For information on all Elsevier publications
visit our website at http://store.elsevier.com/

Editorial & Production: Gemma Tomalin, Richard Berryman, Marise Willis, Michael Nicholls, Ellen Bishop
Cover design: Matthew Limbert

Printed and bound in the United Kingdom

15 16 17 18 19 10 9 8 7 6 5 4 3 2 1

CONTENTS

Preface *vii*

Editor-in-Chief *ix*

Contributors *xi*

Contents of all Volumes *xiii*

4.01 Earthquake Seismology: An Introduction and Overview 1
GC Beroza and H Kanamori

4.02 Seismic Source Theory 51
R Madariaga

4.03 Fracture and Frictional Mechanics: Theory 73
Y Fialko

4.04 Applications of Rate- and State-Dependent Friction to Models of Fault-Slip and Earthquake Occurrence 93
JH Dieterich

4.05 The Mechanics of Frictional Healing and Slip Instability During the Seismic Cycle 111
C Marone and DM Saffer

4.06 Mechanisms for Friction of Rock at Earthquake Slip Rates 139
TE Tullis

4.07 The Role of Fault-Zone Drilling 161
MD Zoback, S Hickman, and W Ellsworth

4.08 Dynamic Shear Rupture in Frictional Interfaces: Speeds, Directionality, and Modes 183
AJ Rosakis, K Xia, G Lykotrafitis, and H Kanamori

4.09 Slip Inversion 215
S Ide

4.10 Fault Interaction, Earthquake Stress Changes, and the Evolution of Seismicity 243
GCP King and MH Devès

4.11 Dynamic Triggering 273
DP Hill and SG Prejean

4.12 Earthquake Hydrology 305
M Manga and C-Y Wang

4.13 Deep Earthquakes 329
H Houston

4.14 Volcanology 101 355
CG Newhall

4.15 Volcano Seismology 389
H Kawakatsu and M Yamamoto

4.16 Interaction of Solid Earth, Atmosphere, and Ionosphere 421
T Tanimoto, K Heki, and J Artru-Lambin

4.17 Episodic Aseismic Slip at Plate Boundaries 445
SY Schwartz

4.18 Global Seismicity: Results from Systematic Waveform Analyses, 1976–2012 467
G Ekström

4.19 Tsunamis 477
K Satake

4.20 Physical Processes That Control Strong Ground Motion 505
JG Anderson

4.21 Paleoseismology 559
L Grant Ludwig

4.22 Historical Seismicity: Archaeoseismology 581
A Nur

4.23 Earthquake Hazard Mitigation: New Directions and Opportunities 593
RM Allen

4.24 Complexity and Earthquakes 627
R Shcherbakov, DL Turcotte, and JB Rundle

PREFACE

The *Treatise on Geophysics* is a comprehensive and in-depth study of the physics of the Earth and of terrestrial-like planets residing both within and outside our solar system. Its breadth and detail of coverage are beyond what any single geophysics text can provide. The first edition of the *Treatise on Geophysics* was published in 2007, nearly a decade ago. Of course, there is much progress in science in that length of time and the field of geophysics has grown rapidly and developed new lines of inquiry. Accordingly, we have brought the *Treatise on Geophysics* up-to-date with this second edition. The new edition will continue to provide students and professionals with fundamental and state-of-the-art discussion of all aspects of geophysics. In this new edition, the reader will find updates to all the chapters contained in the first edition and new chapters that discuss topics missing from the first edition. In a few instances, chapters from the first edition have simply been reprinted in order to avoid gaps in coverage. Chapters that emphasized fundamental physics often required little updating.

A highlight of the second edition is a new volume on Resources in the Near-Surface Earth. The new volume discusses the role of geophysics in the exploitation and conservation of natural resources and the assessment of degradation of natural systems by pollution. The near surface is a zone where humans and natural systems interact. Understanding the effects of human impacts is a challenge particularly for the long term.

There are four new chapters in the volume on Deep Earth Seismology. One of them describes progress in seismic instrumentation in the last 30 years in the era of broadband digital sensor and recording systems. Another one addresses a new approach for the determination of crust and uppermost mantle structure, based on ambient noise cross-correlations. This has revolutionized the field, since it allows the study of regional structure at high resolution in areas where there are no earthquakes. The third new chapter is devoted to recent advances in characterizing the lithosphere–asthenosphere boundary under both oceans and continents. Finally, the last new chapter discusses the morphology of subducted slabs in the Earth's transition zone.

The Mineral Physics volume has new chapters on dynamic (shock-wave) compression methods, and on the mineralogy of the upper mantle and the D″ layer focusing on post-perovskite, and extrasolar super-Earths.

New material in the Geodesy volume expands the temporal resolution and latency of geodetic measurements to high-rate, real-time measurements and increases the spatial resolution to meter size pixels with optical and radar imaging systems. Application areas are extended with new material on the uses of Earth gravity measurements.

The Earthquake Seismology volume has a new chapter on rock mechanics that is a comprehensive account of modern rock mechanics with emphasis on the variation of frictional properties of rocks during an earthquake cycle. Many chapters have considerable updates reflecting the lessons from significant recent earthquakes (e.g., the 2011 Tohoku-Oki earthquake and the 2013 Okhotsk deep earthquake), new observations on slow slips and triggered earthquakes, and developments of new methods (e.g., back projection method). Some chapters are updated (e.g., paleoseismology and strong ground motion) with substantial new material.

Chapters in the Geomagnetism volume have been updated to include recent advances in the acquisition of larger data sets and higher quality data and in the construction of more reliable and sophisticated models. Improvements in data have occurred in areas such as magnetic anomalies, paleointensities, and planetary magnetism. Improvements in models of the present day field include so-called comprehensive models and models that incorporate data assimilation.

The Crustal and Lithosphere Dynamics volume has new contributions on mountain building and landscape evolution.

The Mantle Dynamics volume contains new chapters on the generation of plate tectonics from mantle convection and the core–mantle boundary region.

The Core Dynamics volume features a comprehensive update of this rapidly changing subject, including new chapters on turbulence and waves in the fluid outer core.

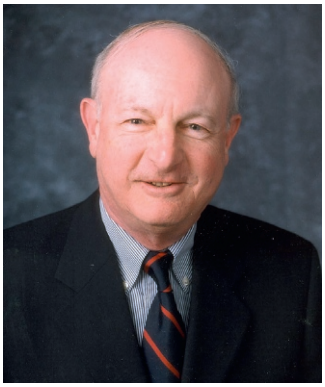
The volume on the Evolution of the Earth presents a new assessment on the role of volatiles in Earth evolution.

There are new chapters in the Physics of Terrestrial Planets and Moons volume on planetary atmospheres, planetary tectonics, extrasolar planets and asteroids and comets. The study of extrasolar planets has had explosive growth in the last decade and new spacecraft missions have significantly enhanced our knowledge of solar system planets and moons.

Finally, this is an opportunity for me to thank and acknowledge the editors of the individual volumes for their hard work, patience, dedication, and commitment in bringing this new edition of the Treatise to fruition. I am also indebted to Gemma Tomalin of Elsevier for her care and attention to the organization and production of the Treatise and to Deepa Arumugam for her careful coordination of the proofs from the copyeditor and typesetter. Of course, the authors of the individual chapters are the ones who are responsible for the excellence of the Treatise. We have all had some ups and downs in the process of producing the second edition of the Treatise but the final product is one that we can all view with pride.

Gerald Schubert
Editor-in-Chief

EDITOR-IN-CHIEF



Gerald Schubert is Distinguished Professor Emeritus in the Department of Earth, Planetary and Space Sciences at the University of California, Los Angeles. His research interests encompass the physics of the interiors and atmospheres of the Earth, the Moon, and the other moons and planets of the solar system. He is co-author with Donald Turcotte of *Geodynamics* (ed. 3, Cambridge University Press, 2014) and with Turcotte and Peter Olson of *Mantle Convection in the Earth and Planets* (Cambridge University Press, 2001). He is the author of over 540 research papers. He has participated in a number of NASA's planetary missions, including Apollo, Pioneer Venus, Magellan, and Galileo, and has been editor and editorial board member of many journals, including *Icarus*, *Journal of Geophysical Research*, *Geophysical Research Letters*, and *Annual Reviews of Earth and Planetary Sciences*. Professor Schubert is a Fellow of the American Geophysical Union and a recipient of the Union's James B. MacElwane medal and the Harry H. Hess medal. He is a member of the National Academy of Sciences and the American Academy of Arts and Sciences.

This page intentionally left blank

CONTRIBUTORS

RM Allen
University of California Berkeley, Berkeley, CA, USA

JG Anderson
University of Nevada, Reno, NV, USA

J Artru-Lambin
Centre National d'Etudes Spatiales, Toulouse, France

GC Beroza
Stanford University, Stanford, CA, USA

MH Devès
Institut de Physique du Globe de Paris, Sorbonne Paris Cit.é, Univ. Paris Diderot, Paris, France

JH Dieterich
University of California, Riverside, CA, USA

G Ekström
Columbia University, Palisades, NY, USA

W Ellsworth
US Geological Survey, Menlo Park, CA, USA

Y Fialko
University of California San Diego, La Jolla, CA, USA

L Grant Ludwig
University of California, Irvine, CA, USA

K Heki
Hokkaido University, Sapporo, Japan

S Hickman
US Geological Survey, Menlo Park, CA, USA

DP Hill
US Geological Survey, Volcano Science Center, Menlo Park, CA, USA

H Houston
University of Washington, Seattle, WA, USA

S Ide
University of Tokyo, Tokyo, Japan

H Kanamori
California Institute of Technology, Pasadena, CA, USA

H Kawakatsu
The University of Tokyo, Tokyo, Japan

GCP King
Institut de Physique du Globe de Paris, Sorbonne Paris Cit.é, Univ. Paris Diderot, Paris, France

G Lykotrafitis
Massachusetts Institute of Technology, Cambridge, MA, USA

R Madariaga
Laboratoire de Géologie Ecole Normale Supérieure, Paris, France

M Manga
University of California, Berkeley, CA, USA

C Marone
The Pennsylvania State University, University Park, PA, USA

CG Newhall
Mirisbiris Garden and Nature Center, Sto. Domingo, Albay, Philippines

A Nur
Stanford University, Stanford, CA, USA

SG Prejean
US Geological Survey, Alaska Volcano Observatory, Anchorage, AK, USA

AJ Rosakis
California Institute of Technology, Pasadena, CA, USA

JB Rundle
University of California, Davis, CA, USA

DM Saffer
The Pennsylvania State University, University Park, PA, USA

K Satake
University of Tokyo, Tokyo, Japan

SY Schwartz
University of California Santa Cruz, Santa Cruz, CA, USA

R Shcherbakov
Western University, London, ON, Canada

T Tanimoto
University of California, Santa Barbara, CA, USA

TE Tullis
Brown University, Providence, RI, USA

DL Turcotte
University of California, Davis, CA, USA

C-Y Wang
University of California, Berkeley, CA, USA

K Xia
University of Toronto, Toronto, ON, Canada

M Yamamoto
Tohoku University, Sendai, Miyagi, Japan

MD Zoback
Stanford University, Stanford, CA, USA

4.01 Earthquake Seismology: An Introduction and Overview

GC Beroza, Stanford University, Stanford, CA, USA
H Kanamori, California Institute of Technology, Pasadena, CA, USA

© 2015 Elsevier B.V. All rights reserved.

4.01.1	**Introduction**	2
4.01.2	**Seismicity**	3
4.01.2.1	Earthquake Size	3
4.01.2.2	Earthquakes in the Context of Plate Tectonics	5
4.01.2.2.1	Transcurrent plate boundaries	6
4.01.2.2.2	Divergent plate boundaries	7
4.01.2.2.3	Convergent plate boundaries	7
4.01.2.2.4	Intraplate earthquakes	11
4.01.2.2.5	Hot spot volcanism	13
4.01.2.3	The Largest and Deadliest Earthquakes	13
4.01.2.4	Historic and Prehistoric Earthquakes	15
4.01.2.5	Earthquake Size Distribution	15
4.01.2.6	Earthquake Location	16
4.01.3	**The Earthquake Source**	17
4.01.3.1	Point-Source Parameters	17
4.01.3.2	Sense of Faulting from First Motions	18
4.01.3.3	Moment Tensor Representation	19
4.01.3.4	Seismic Energy	19
4.01.3.5	Extended-Source Models of Earthquakes	20
4.01.3.5.1	Kinematic source models	20
4.01.3.5.2	Dynamic source models	20
4.01.3.6	Volcano Seismology	22
4.01.4	**Slip Behavior**	24
4.01.4.1	Brittle Failure	24
4.01.4.2	Creep	25
4.01.4.3	Aseismic Transients and Tremor	27
4.01.5	**Physics of the Earthquake Source**	30
4.01.5.1	Friction	30
4.01.5.2	Energy Budget	31
4.01.5.3	Microscopic Processes	31
4.01.5.3.1	Hydrologic processes	32
4.01.5.3.2	Melting and flash heating	32
4.01.5.3.3	Thermal pressurization	33
4.01.5.3.4	Silica gel formation	33
4.01.5.4	Fault-Zone Structure	34
4.01.5.5	Borehole Observations, Fault-Zone Drilling, and Seismicity in Mines	34
4.01.5.6	Earthquakes as a Complex System	35
4.01.6	**State of Stress, Nucleation, and Triggering**	35
4.01.6.1	State of Stress	35
4.01.6.2	Earthquake Nucleation and Short-Term Earthquake Prediction	37
4.01.6.3	Earthquake Forecasting	38
4.01.6.4	Static Stress Triggering	39
4.01.6.5	Dynamic Triggering	40
4.01.6.6	Temporal Distribution of Earthquakes	40
4.01.7	**Associated Problems**	40
4.01.7.1	Strong Motion Prediction	40
4.01.7.2	Tsunamis	43
4.01.7.3	Test Ban Treaty Verification	43
4.01.7.4	Solid Earth–Atmospheric Coupling	43
4.01.8	**Earthquake Risk Mitigation**	44
4.01.9	**Conclusions**	44
References		46

 http://dx.doi.org/10.1016/B978-0-444-53802-4.00069-5

4.01.1 Introduction

In general usage, the term 'earthquake' describes a sudden shaking of the ground. Earth scientists, however, typically use the word 'earthquake' differently – to describe the *source* of seismic waves, which is nearly always sudden shear slip on a fault within the Earth. In this chapter, we follow the scientific usage of the term and focus our review on how earthquakes are studied using the motion of the ground remote from the earthquake source itself, that is, by interpreting the same shaking that most people consider to be 'the earthquake.' The field defined by the use of seismic waves to understand earthquakes is known as earthquake seismology (Figure 1).

The nature of the earthquakes makes them intrinsically difficult to study. Different aspects of the earthquake process span a tremendous range in length scales – all the way from the size of individual mineral grains to the size of the largest plates. They span a tremendous range in timescales as well. The smallest microearthquakes rupture faults for only a small fraction of a second, and the duration of even the very largest earthquakes can be measured in 100 s of seconds. Compare this with the length of strain accumulation in the earthquake cycle, which can be measured in decades, centuries, and even millennia in regions of slow strain rate. The evolution of fault systems spans longer times still, since that can require the action of thousands of earthquakes. At different physical dimensions or temporal scales, different physical mechanisms may become important or perhaps negligible. Earthquakes occur in geologically, and hence physically, complicated environments. The behavior of earthquakes has been held up as a type example of a complex natural system. The sudden transformation of faults from being locked, or perhaps slipping quasistatically, to slipping unstably at large slip speeds, as is nearly universally observed for earthquakes, also makes them a challenging physical system to understand.

Figure 1 Earthquakes are due to slip on faults within the Earth. In large earthquakes, fault slip can reach the Earth's surface. Photo shows surface rupture of the 1906 San Francisco earthquake offset this fence horizontally by ~8.5 ft. (Plate 1-B, US Geological Survey Bulletin 324 – from http://libraryphoto.cr.usgs.gov/, San Francisco Earthquake, plate 41.)

Despite these challenges, seismologists have made tremendous progress in understanding many aspects of earthquakes – elucidating their mechanisms based on the radiated seismic wave field, determining where they occur and the deep structure of faults with great precision, documenting the frequency and the regularity (or irregularity) with which they occur (and recur) over the long term, gaining insight into the ways in which they interact with one another, and so on. Yet, the obvious goal of short-term prediction of earthquakes, that is, specifying the time, location, and size of future significant earthquakes on a timescale shorter than decades, remains elusive. Earthquakes are different in this sense from nearly all other deadly natural hazards such as hurricanes, floods, tornadoes, and even volcanic eruptions, which to varying degrees are predictable over a timescale of hours to days. The worst earthquakes rank at the very top of known disasters. The deadliest known earthquake killed over half a million people in a matter of minutes, a level of sudden destruction that no other catastrophe in recorded history – either natural or human-made – has attained.

Our inability to predict earthquakes is one reason they cause such apprehension. This lack of warning is compounded by the fact that they strike so abruptly. No one can see an earthquake coming and it is only a matter of seconds from the initial perception of the first-arriving waves of a large earthquake before dangerous strong ground motion begins. Moreover, large, damaging earthquakes occur infrequently (fortunately) at any given point on the Earth relative to a human life span. This means that most people who experience a major earthquake are doing so for the first time, so the experience is unfamiliar. Finally, there is something fundamentally unsettling about the movement of the 'solid' Earth.

The unpredictability, sudden onset, and their unfamiliarity make earthquakes a uniquely terrifying phenomenon. As testament to this, other extreme and catastrophic events in the affairs of humankind – if they are devastating enough – are described with the simile 'like an earthquake.' The point of origin of an extreme event of any kind is often described as 'the epicenter,' a term borrowed from seismology.

The unpredictability of earthquakes also renders them difficult to study. Since we do not know where, and particularly when, large earthquakes will strike, collecting data on earthquakes has to be approached passively. Seismologists deploy instruments to measure seismic waves where we expect earthquakes to occur and then wait for nature to carry out the experiment. The wait can last decades, or more, for a large earthquake. Inevitably, with finite budgets and finite patience, this leads to seismic monitoring instruments being widely, and hence too thinly, dispersed in an attempt to gather data from at least *some* earthquakes wherever and whenever they might occur.

Finally, the combination of unpredictability, sudden onset, long intervals between events, and unfamiliarity means that the risk created by earthquake hazards is extremely difficult for both policymakers and the general public to manage. Because earthquakes are not predicted and occur infrequently relative to other hazards, it is understandably tempting for

governments and individuals to focus on the many immediate and predictable problems that impact society more frequently. The unpredictability and sudden onset of earthquakes, however, mean that once an earthquake begins, it is too late to do much more than duck, cover, and hold on.

4.01.2 Seismicity

4.01.2.1 Earthquake Size

Perhaps, the most familiar earthquake source parameter is magnitude. Earthquake magnitude and the characteristics of earthquake behavior that it is used to define, such as Gutenberg–Richter statistics (discussed further in Section 4.01.2.5), are purely empirical observations in the sense that they rely only on measurements of seismic waves as recorded on seismographs and do not require much in the way of assumptions about earthquake source physics. It is impossible to completely represent a complex physical process like an earthquake with a single number that measures its size. Nevertheless, several definitions of earthquake size have proven extremely useful for reaching a better understanding of earthquakes.

Earthquake size is traditionally measured by one of various magnitude scales. M_L, the local magnitude scale, was devised by Richter in the early 1930s (Richter, 1935). He was cataloging data from the Southern California seismic network, and although locations, depths, and origin times for many earthquakes were available, there was no measurement of earthquake size. So he invented one, using the relation

$$M_L = \log(A) - \log(A_0)$$

where A is the measured amplitude of the seismic trace on a standardized Wood–Anderson seismograph (Figure 2) at a distance of 100 km and A_0 is the amplitude of a reference earthquake with $M_L = 0.0$. The particular definition of the zero level is that an earthquake of magnitude 0 has an amplitude of 0.000001 m at a distance of 100 km. Thus, an earthquake of magnitude 3.0 has an amplitude of 1 mm at the same distance and a magnitude 7.0 earthquake would have an amplitude of 10 m. Wood–Anderson instruments record on photographic paper and saturate at an amplitude of about 20 cm.

Until recently, quoted magnitudes for local earthquakes were often correctly described as the Richter magnitude. The old

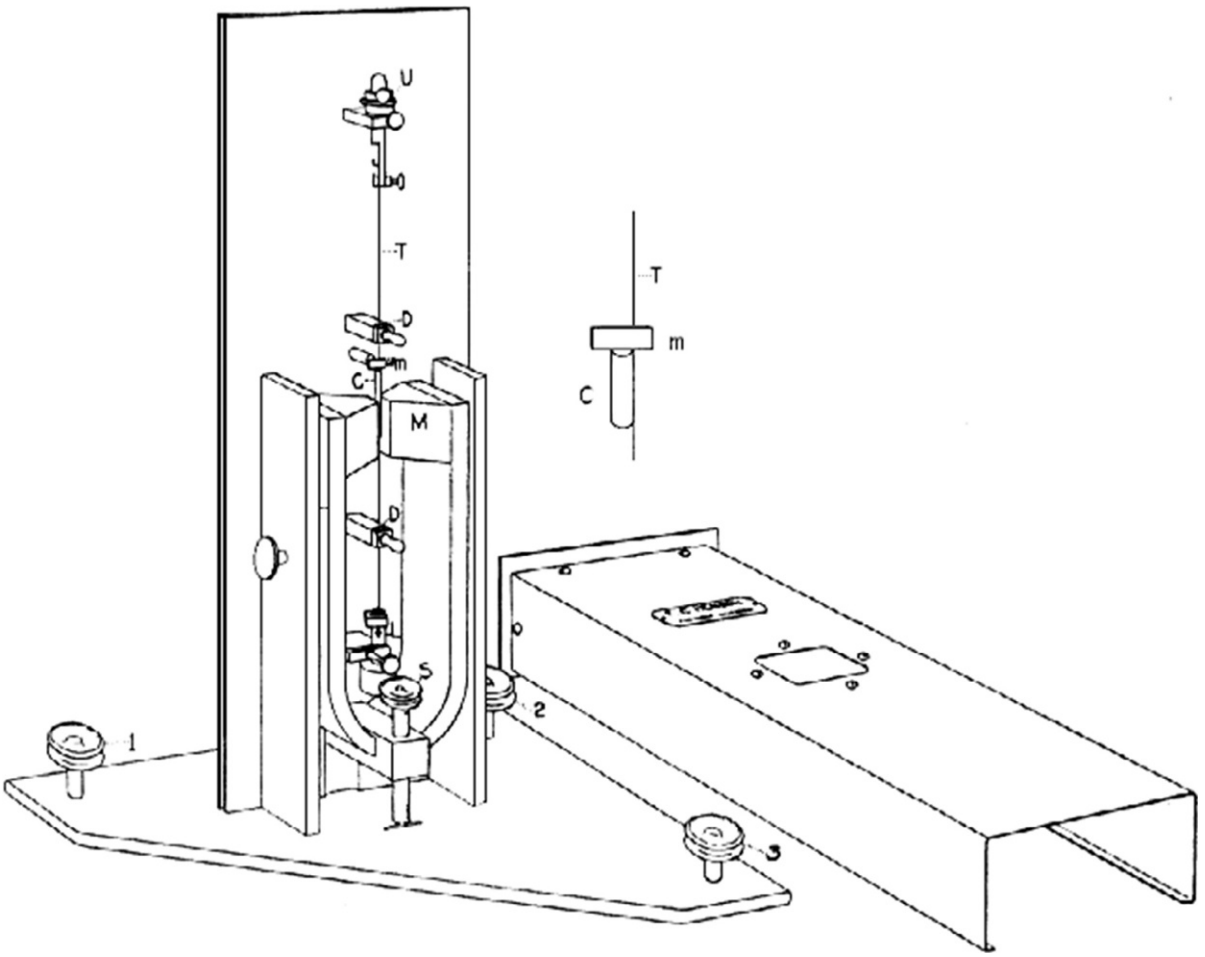


Figure 2 Diagram of a Wood–Anderson torsion seismometer, which is used to define M_L. Ground motion causes rotation of an inertial mass, C, which is attached to a thin wire under high tension, T. Seismograms are recorded on photographic paper using a beam of light reflected off a mirror, M, that is mounted on the inertial mass. Reproduced from http://www.data.scec.org/Module/s3inset3.html.

Wood–Anderson instruments, though state of the art when the magnitude scale was developed, are not now used. Instead, if one wants to calculate the local magnitude, data from modern instruments must be degraded to mimic recordings from the old instruments (Figure 3). Although the magnitude scale has proven extremely useful for routine cataloging of earthquake behavior, there are a several shortcomings of M_L, among them:

(1) If the station is not 100 km from the earthquake, one must correct for propagation effects – this is difficult because at different distances, different wave types have the largest amplitude.
(2) Excitation of waves can vary strongly with depth and this must be corrected for to compare the size of earthquakes at different depths.
(3) Earthquakes have complex radiation patterns so that the azimuthal direction of the seismograph from the earthquake strongly affects the amplitude of the recorded waves.
(4) The Wood–Anderson instruments respond primarily to ground motion at periods less than about 1 s. Once the size of an earthquake approaches magnitude 7.0, 1 s waves start to saturate in amplitude due to interference effects arising from the finite size and duration of the fault. This occurs even if the seismograph is operated at low gain and the waves are all on scale. Beyond $M_L \sim 7$, the local magnitude is no longer a reliable measure of earthquake size.

To some extent, difficulties (1)–(3) can be corrected for empirically and by averaging over many stations. Nevertheless, M_L is frequently uncertain by as much as 0.25 magnitude units, and it does not extend to earthquakes recorded at distances in excess of 300 km. Hence, it is not suitable for global earthquake monitoring. To remedy this, and to address difficulty (4), seismologists invented a variety of different magnitude scales. Usually, it is one of these other magnitude scales that is used to characterize the size of an earthquake.

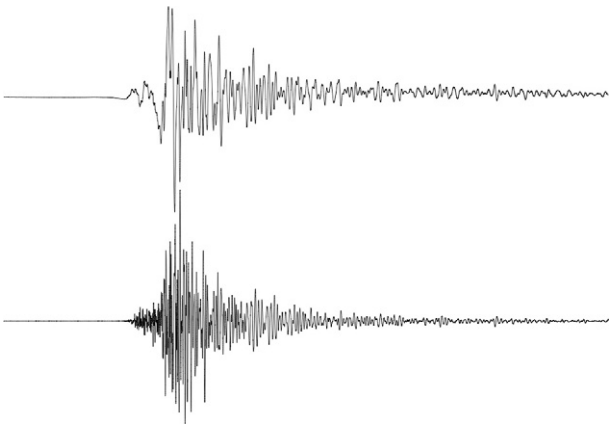


Figure 3 Upper seismogram shows data for the 28 September 2004 Parkfield, California, earthquake from a modern broadband instrument at station JRSC (Jasper Ridge, Stanford, California), and lower seismogram shows the same ground motion as it would have been recorded on a Wood–Anderson instrument (lower). Time increases from left to right and a total of 600 s is shown. Vertical axis shows north–south component of displacement. Maximum peak-to-peak displacement of the broadband record is ~0.4 cm. At frequencies <1 Hz, the Wood–Anderson instrument is flat to acceleration, which means that the displacement response decays rapidly. The lower seismogram is multiplied by a factor of 20 for comparison. Standard magnification for these instruments was 2800, which would have slightly exceeded the dynamic range of the photographic recording system for this earthquake.

m_b, the body wave magnitude scale, is defined using the initial P-waves from earthquakes that occur very far away so that it can be used to compare earthquakes all over the world. It has been adjusted to agree approximately with the Richter magnitude at short distances. Since it uses short-period waves, it also saturates above magnitude 7.0. The application of the body wave magnitude scale has varied with time – particularly in whether or not the maximum amplitude of the entire P-wave train, rather than just the first few cycles, is measured. Prior to 1964, it was determined using the formula

$$m_b = \log(A/T) + Q$$

where A is the amplitude, T is the period, and Q is a factor that corrects for distance and depth. Unlike the local magnitude scale, the amplitude measurement for the body wave magnitude represents the ground motion amplitude after correcting for instrument response, rather than the amplitude as recorded on the instrument. With the advent of the Worldwide Standardized Seismographic Network (WWSSN), m_b was defined using the high-frequency channel of those instruments, for which the response is strongly peaked at a period of 1 s.

M_s, the surface wave magnitude scale, is a measure of the amplitude of surface waves (Rayleigh waves) at a period of 20 s as measured very far from the event. This magnitude is often quoted by the press as the Richter magnitude. The internationally adopted definition (Vanek et al., 1962) of the surface wave magnitude scale is

$$M_s = \log(A/T)_{max} + 1.66\log\Delta + 3.3$$

where A is the peak amplitude of vertical ground motion (as for the body wave magnitude, it is corrected for instrument response) and Δ is the angular distance from the earthquake to the seismograph. This definition of the surface wave magnitude closely resembles the original definition of Gutenberg and Richter (1936), which used a fixed period of 20 s. M_s is usually a better estimate of earthquake size, especially for larger earthquakes, than either m_b or M_L. It only saturates for 'great' earthquakes when M_s is much over eight. For intermediate and deep earthquakes, however, it does a poor job of representing earthquake size because such events do not strongly excite surface waves.

The magnitude scales described earlier do not tie magnitude to a physical parameter. An obvious physical parameter of interest that is strongly linked to the strength of an earthquake is the total amount of energy radiated in the form of seismic waves. This is often referred to as the radiated seismic energy, or the seismic energy for short. Gutenberg and Richter (1942, 1956) attempted to relate the seismic energy to the surface wave magnitude. For the seismic energy, E_s, expressed in Joules, they found

$$\log E_s = 1.5M_s + 4.8$$

This relationship assumes proportionality between the energy radiated as Rayleigh waves at ~20 s period and the

entire radiated field. As with the surface wave magnitude scale, this is a reasonable assumption until M_s approaches eight. For larger earthquake, the effects of fault finiteness start to become important for 20 s Rayleigh waves and the surface wave magnitude is no longer representative of the total radiated energy. Nevertheless, this relationship shows that the radiated energy increases very rapidly with magnitude. A unit increase in magnitude corresponds to a factor of 32 ($10^{1.5}$) increase in the seismic energy.

M_0, the seismic moment, is another physical parameter related to the size of an earthquake and has become the parameter that seismologists most often use to describe the size of an earthquake. The seismic moment is a measure of the size of the equivalent force system needed to generate the waves that the earthquake faulting generated. M_0 was first calculated by Aki (1966) in his study of the Niigata, Japan, earthquake. It is a more robust estimate of the size because it is done, in theory, at infinite period and radiation pattern effects due to the orientation of faulting are explicitly accounted for when it is calculated. Another advantage it has is that it can be applied to aseismic transients, which by definition do not radiate detectable seismic waves. Seismic moment does not saturate and is directly proportional to the amount of slip across a fault times the fault area. If the faulted area is A, the shear modulus is μ, and the average slip over the faulted area is s, then the seismic moment, M_0, is

$$M_0 = \mu A s$$

The MKS unit for seismic moment is Newton meters. Earlier work in seismology typically used dyne cm, the cgs unit, to express seismic moment.

M_w, the moment magnitude scale, is a magnitude scale based directly on the measurement of the seismic moment so it does not saturate. M_w was constructed to agree with the M_L and M_s scales before they saturate but retains the same relationship to seismic moment thereafter (Hanks and Kanamori, 1979). Assuming the seismic moment is expressed in Nm, the moment magnitude relation can be written

$$\log M_0 = 1.5 M_w + 9.05$$

and hence, the moment magnitude scale is determined from the seismic moment by the relation

$$M_w = 2/3 \log M_0 - 6.03$$

A unit increase in M_w does *not* mean a factor of 10 increase in earthquake size. Rather, the increase in size of the earthquake, as measured by M_0, is about factor of 32 ($10^{1.5}$) for each magnitude unit.

As is apparent from all of the definitions mentioned previously, magnitude scales do not vary from 1 to 10, but are open-ended. In practice, even M_w is limited on the high end both by the strength of the rocks that earthquakes occur in and by the size of the fault. On the low end, M_w is limited by the size of the smallest faults and/or fault patch that can slip unstably. Earthquakes that are recorded by sensitive borehole instruments or *in situ* in deep mines routinely have negative magnitudes. There are many other magnitude scales, most of which are tailored to suit either the practicalities or different goals of earthquake, volcano, and nuclear explosion monitoring. These include magnitude scales derived from measurements as diverse as the duration of the coda of scattered waves following an earthquake, the amplitude of guided waves in the continental crust, and the amplitude of the tsunami that an earthquake generates.

Instrumental seismology began in the last decade of the nineteenth century; however, many earthquakes of interest happened before then. For such earthquakes, seismologists must resort to more subjective and indirect estimates of earthquake magnitude. The most common approach to determining the magnitude of preinstrumental earthquakes is through the interpretation of seismic intensity observations.

Intensity is different from magnitude in that it is not a measure of the intrinsic size of an earthquake; rather, it is a local measure of the strength of shaking as perceived by people or as it affects buildings or natural features. There are many different intensity scales, but the most widely used is probably the modified Mercalli intensity (MMI) scale (Richter, 1958). Traditionally, intensities are denoted with Roman numerals, and the MMI scale ranges from I to XII. **Table 1** contains the MMI scale as abridged by Bolt (1993).

Earthquake magnitude can be determined from the intensity distribution in several ways. Some methods use the area encompassed by intensities of a certain level (Toppozada, 1975), whereas others use individual intensity observations directly (Bakun and Wentworth, 1997). Toppozada et al. (1981) used intensity distributions to estimate the magnitudes of California earthquakes from the preinstrumental period. Because the stronger intensities are direct measures of damage in earthquakes, there have been attempts to relate intensities to measures of ground motion, like peak ground acceleration (PGA), that are thought to correlate strongly with damage (Gutenberg and Richter, 1956). More recently, instrumental recordings of ground motion have been used to estimate instrumental intensities. Wald et al. (1999) found the following relation between PGA and the MMI level, I_{mm}:

$$I_{mm} = 3.66 \log(\mathrm{PGA}) - 1.66 \; (\sigma = 1.08)$$

Intensity data have also been used as a proxy to study strong ground motion from significant earthquakes where such data are lacking, such as for the 1906 San Francisco, California, earthquake (Boatwright and Bundock, 2005). The intensity map of the 1886 Charleston, South Carolina, earthquake is shown in **Figure 4**.

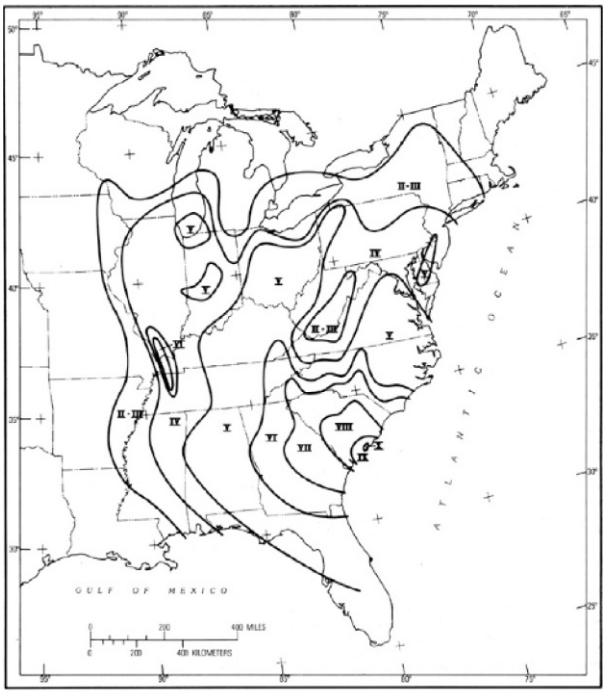
4.01.2.2 Earthquakes in the Context of Plate Tectonics

Scientists have long recognized that some portions of the Earth are characterized by frequent earthquake activity, while others have essentially no earthquake activity whatsoever. **Figure 5** shows the first global map of earthquake activity (Mallet and Mallett, 1858), which shows a remarkably accurate view of global seismicity, particularly since it predated the invention of the seismograph by decades.

In 1954, Beno Gutenberg and Charles Richter published the second edition of *Seismicity of the Earth*. To this day, their book provides a remarkably complete description of the locations, depth, and size of earthquakes, as well as the distribution and types of volcanoes, worldwide (**Figure 6**). What the book lacks

Table 1 The modified Mercalli intensity scale

I	Not felt except by a very few under especially favorable circumstances
II	Felt only by a few persons at rest, especially on upper floors of buildings. Delicately suspended objects may swing
III	Felt quite noticeably indoors, especially on upper floors of buildings, but many people do not recognize it as an earthquake. Standing automobiles may rock slightly. Vibration like passing of truck
IV	During the day, felt indoors by many, outdoors by few. At night, some awakened. Dishes, windows, doors disturbed; walls make creaking sound. Sensation like heavy truck striking building. Standing automobiles rocked noticeably
V	Felt by nearly everyone, many awakened. Some dishes, windows, and so on broken; cracked plaster in a few places; unstable objects overturned. Disturbances of trees, poles, and other tall objects sometimes noticed. Pendulum clocks may stop
VI	Felt by all, many frightened and run outdoors. Some heavy furniture moved; a few instances of fallen plaster and damaged chimneys. Damage slight
VII	Everybody runs outdoors. Damage negligible in buildings of good design and construction, slight to moderate in well-built ordinary structures, considerable in poorly built or badly designed structures, some chimneys broken. Noticed by persons driving cars
VIII	Damage slight in specially designed structures, considerable in ordinary substantial buildings with partial collapse, great in poorly built structures. Panel walls thrown out of frame structures. Fall of chimneys, factory stack, columns, monuments, and walls. Heavy furniture overturned. Sand and mud ejected in small amounts. Changes in well water. Persons driving cars disturbed
IX	Damage considerable in specially designed structures; well-designed frame structures thrown out of plumb; great in substantial buildings, with partial collapse. Buildings shifted off foundations. Ground cracked conspicuously. Underground pipes broken
X	Some well-built wooden structures destroyed, most masonry and frame structures destroyed with foundations, ground badly cracked. Rails bent. Landslides considerable from riverbanks and steep slopes. Shifted sand and mud. Water splashed, slopped over banks
XI	Few (masonry), if any, structures remain standing. Bridges destroyed. Broad fissures in ground. Underground pipelines completely out of service. Earth slumps and land slips in soft ground. Rails bent greatly
XII	Damage total. Waves seen on ground surface. Lines of sight and level distorted. Objects thrown into the air



Isoseismal Map of Sept. 1, 1886 South Carolina Earthquake Magnitude 7.3

Figure 4 Isoseismal map for the 1886 Charleston, South Carolina, earthquake. The lines enclosing areas of equal intensity are referred to as 'isoseismal' lines. Reproduced from http://earthquake.usgs.gov/learning/teachers/IsoseismalMaps.pdf.

is a unified model to tie all those observations together, explaining the where and the why of the Earth's earthquake and volcanic activity. The theory of plate tectonics, which was established within a decade following publication of that second edition, provides that context. *Seismicity of the Earth* underscores both just how much was known about the systematics of earthquake and volcanic activity before the theory of plate tectonics was formulated and what a wide range of phenomena plate tectonics explains.

Earthquakes delineate plate boundaries, the vast majority of earthquakes occur at plate boundaries (Figures 7 and 8), and the type of boundary – transcurrent, divergent, and convergent – exerts a profound effect on the nature of earthquake activity on it. Thus, it is natural to organize a discussion of earthquakes and volcanoes by plate boundary type. In this section, we are only able to offer a cursory review of the systematics of earthquakes and volcanoes.

4.01.2.2.1 Transcurrent plate boundaries

Most transcurrent plate boundary earthquakes, in which plates slide past one another horizontally on strike-slip faults, occur on oceanic transform faults deep under the ocean surface. Such oceanic transform fault earthquakes do not pose much hazard to humankind. That and their remoteness from seismic monitoring instruments render them difficult to study; however, some transform plate boundaries traverse continental crust where they do pose a significant hazard and are more easily monitored. Transform fault earthquakes, whether in the oceans or on continents, are limited in their size by the depth extent of the seismogenic crust, which is at most a few tens of kilometers, and the fact that transform faults tend to cut through the Earth's crust at angles very close to vertical.

Despite this limitation on size, continental transform fault earthquakes that occur in places such as Turkey, New Zealand, and California can be devastating and include some of the most infamous earthquakes in history. When they occur in continental crust, the shallow extent of the seismogenic zone means that waves are generated everywhere very close to the Earth's surface and hence have little opportunity to spread out or attenuate, and thus reduce in amplitude, before impacting the built environment. Although strike-slip deformation

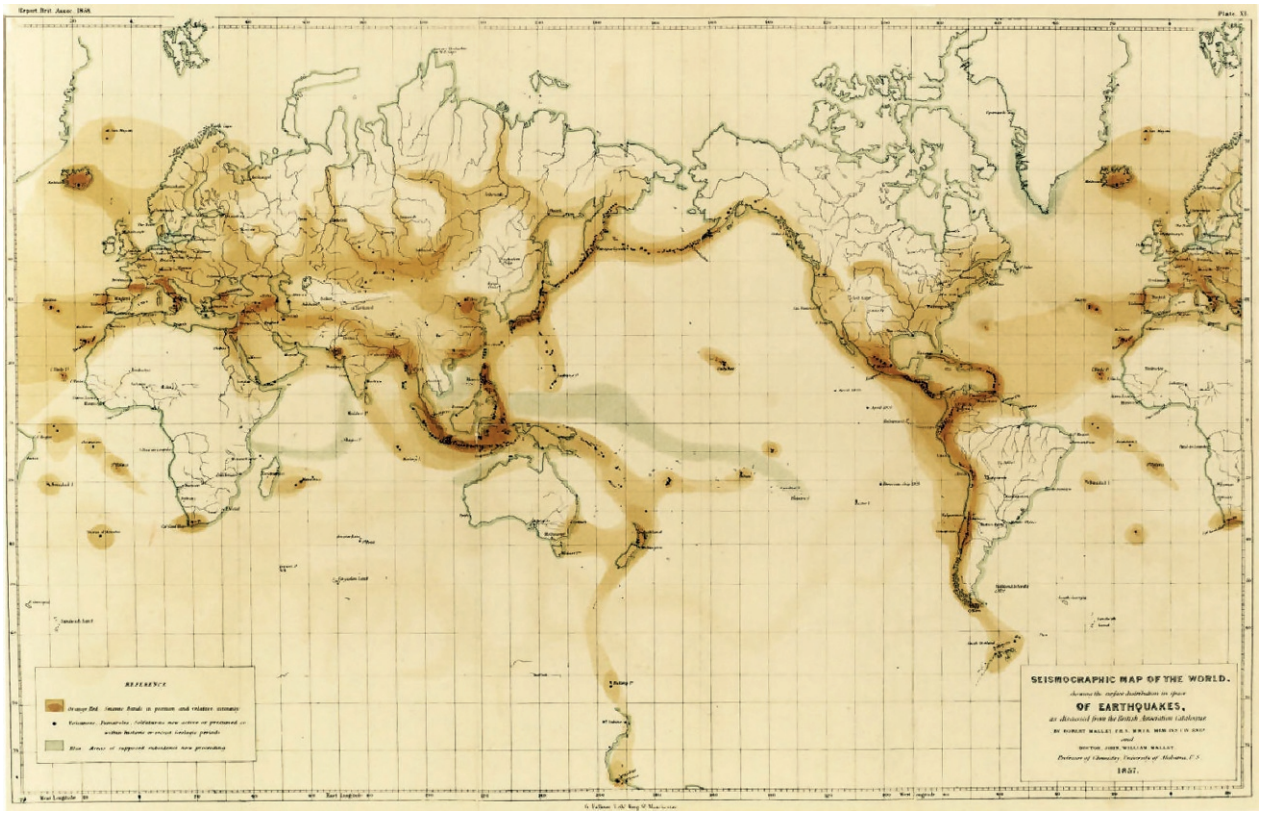


Figure 5 Global distribution of earthquakes as represented by Mallett and Mallett (1858). The principal zones of seismicity are well represented, with the exception of the mid-ocean ridge system, for which the earthquakes were too small and too remote from land to detect. This map preceded instrumental seismology and was based entirely on felt reports of earthquakes.

predominates at transform plate boundaries, such plate boundaries are not perfectly parallel to the direction of relative plate motion. For this reason, there are often significant reverse and normal faulting earthquakes at such boundaries too, particularly for continental transform boundaries, which typically involve much more broadly distributed fault systems than oceanic transforms.

4.01.2.2.2 Divergent plate boundaries

Earthquakes at plate boundaries where plates diverge from one another on normal faults have the least societal impact of any type of plate boundary earthquakes. Most such plate boundaries occur as mid-ocean ridges situated several kilometers beneath the ocean surface. Exceptions to this include Iceland and the Galápagos Islands, where the effects of mid-ocean ridges and hot spots are superimposed, and East Africa, where the divergent plate boundary is not yet mature enough to have separated sufficiently for an ocean to form. Divergent plate boundaries are characterized by effusive basaltic volcanism through which new crust is formed.

A great deal of the deformation at these plate boundaries appears to take place aseismically. The largest normal faulting earthquakes that occur on extensional plate boundaries are small relative to the largest earthquakes at other plate boundaries owing to the fact that the high temperatures in newly formed crust limit the depth extent of the seismogenic crust. Normal faulting predominates at extensional plate boundaries, but extensional rifts are often irregular, and it is not uncommon for strike-slip earthquakes to occur in such settings across local transform faults within a rift. In some cases, these earthquakes are significant.

Mid-ocean ridge volcanism for the most part occurs under several kilometers of ocean. Hence, until the 1950s, the extent and systematics of the mid-ocean ridge system were unknown. Nevertheless, as the source of new oceanic crust, it is a key element of plate tectonics, and the creation of oceanic crust at mid-ocean ridges is *the* key behavior that distinguishes the theory of plate tectonics from the theory of continental drift. The mid-ocean ridge volcanoes form the longest mountain chain on Earth – some 40 000 km in total extent. The hydrothermal plumbing of these magmatic systems is complex, depends on the spreading rate, and gives rise to a range of unusual hydrothermal deep-sea vents. Although eruptions on mid-ocean ridge systems have been inferred from signals detected on hydrophone arrays, they have not yet been directly observed.

4.01.2.2.3 Convergent plate boundaries

By almost any measure, earthquakes at convergent plate boundaries dominate the seismicity of the Earth. At such plate boundaries, all factors work towards the generation of frequent large earthquakes. Relatively cool crust is subducted into the Earth, resulting in reduced temperatures at depth, which increases the depth extent of the seismogenic zone.

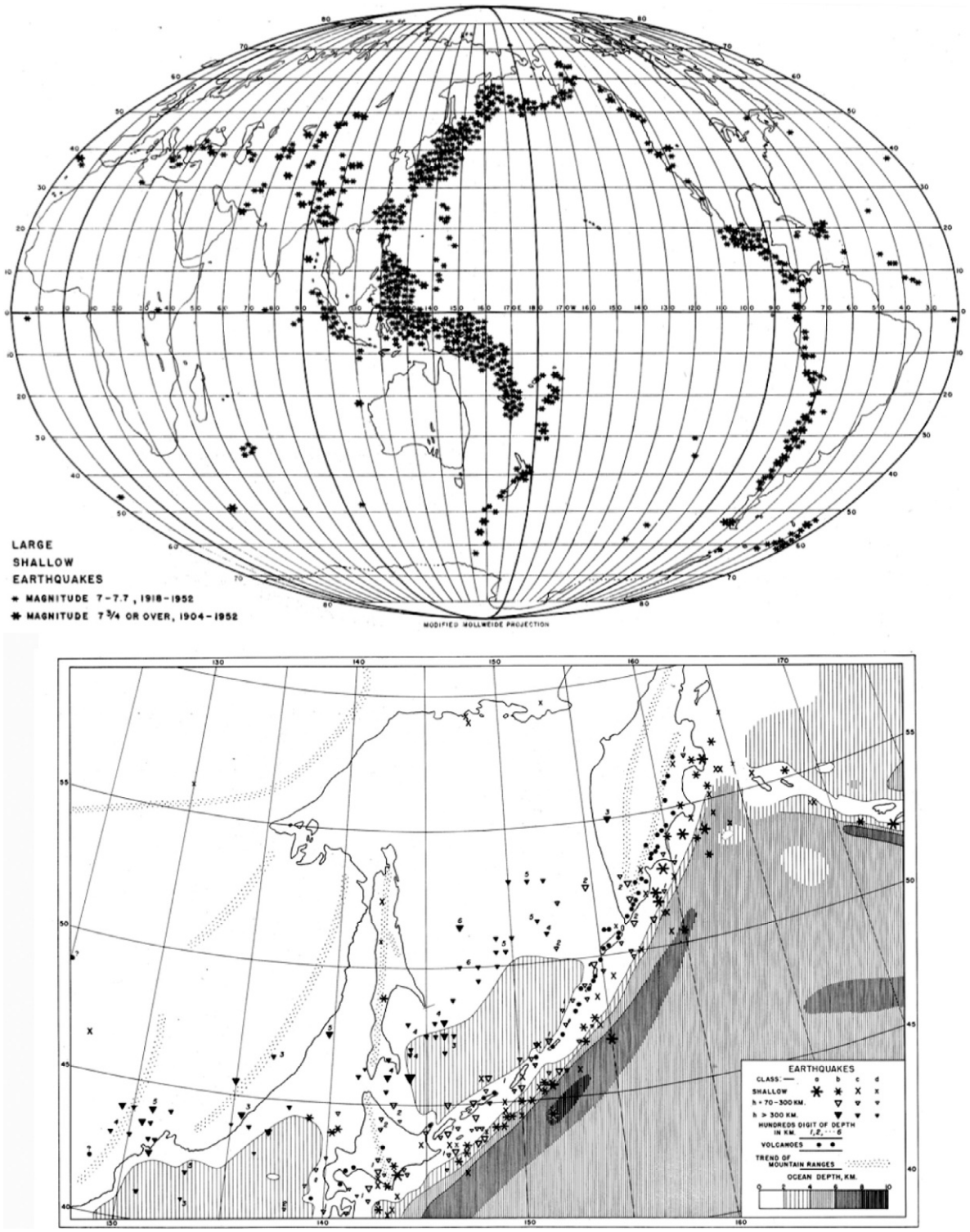


Figure 6 (a) Seismicity map showing worldwide distribution of shallow earthquakes. (b) Close-up of subduction zone seismicity including Honshu, the Kuril Islands, and Kamchatka. The principal features including trench offshore, increasing depth of earthquakes to the northwest, and volcanoes were subsequently explained by the theory of plate tectonics. Reproduced from Gutenberg B and Richter CF (1954) Seismicity of the Earth and Associated Phenomenon, 2nd edn., Princeton: Princeton University Press.

Moreover, unlike other plate boundaries, the plate interface of some subduction zones traverses the seismogenic zone at a very shallow angle – in places $< 10^{\circ}$ from horizontal – meaning that the seismogenic width of the plate boundary in places like Sumatra, Alaska, and Chile can be 100 s of kilometers across. As a result, the great majority of large earthquakes, and nearly all earthquakes larger than magnitude 8, occur at convergent plate boundaries with reverse faulting mechanisms. The very largest

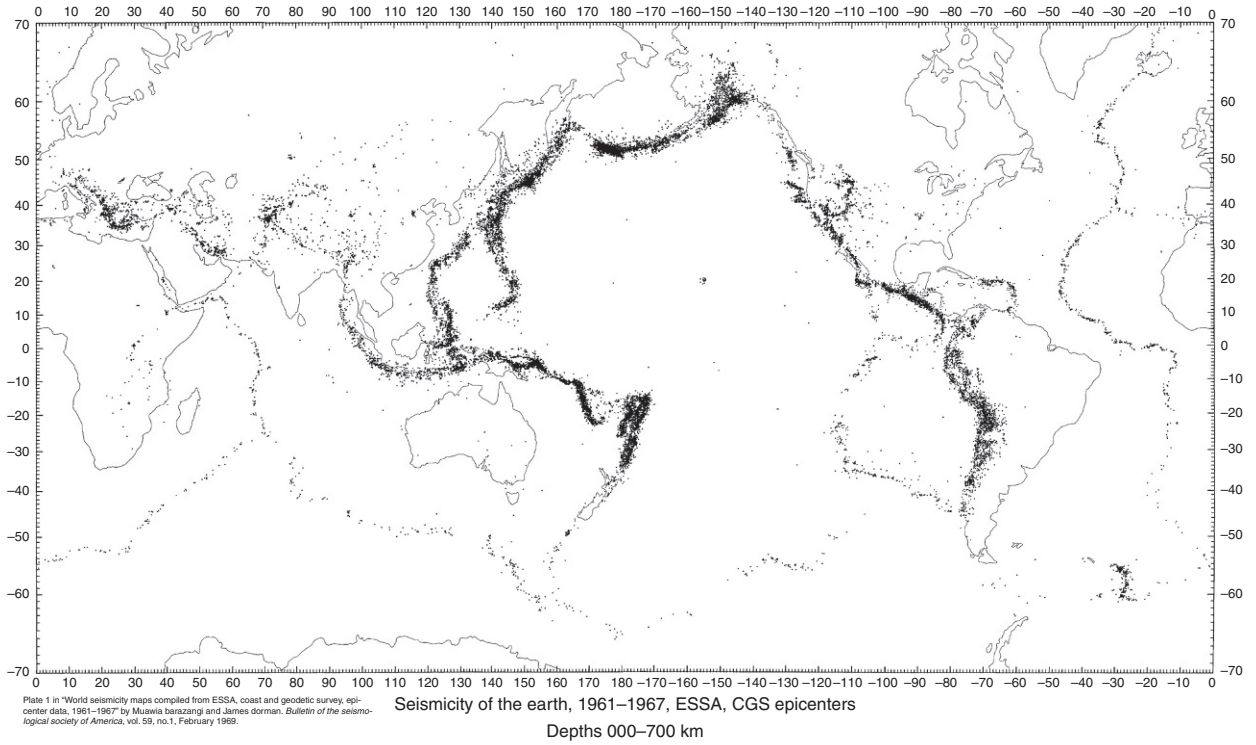


Plate 1 in "World seismicity maps compiled from ESSA, coast and geodetic survey, epicenter data, 1961–1967" by Muawia barazangi and James dorman. Bulletin of the seismological society of America, vol. 59, no.1, February 1969.
Seismicity of the earth, 1961–1967, ESSA, CGS epicenters
Depths 000–700 km

Figure 7 Global seismicity as shown by Barazangi and Dorman (1969) from 1961 to 1967, which comprises 29 553 earthquakes of magnitude >4.0. The narrowbands of seismicity running through the Earth's oceans are readily apparent in this figure.

recorded earthquakes are all of this type, including the M_w 9.5 1960 Chile earthquake, the M_w 9.3 Alaska earthquake of 1964, the M_w 9.2 Sumatra earthquake of 2004, the M_w 9.0 Kamchatka earthquake of 1952, and the M_w 9.0 Tohoku-Oki earthquake of 2011. These five earthquakes dominate the total seismic activity of the Earth over the past 100 years. Of these largest recorded earthquakes, the Sumatra earthquake caused the most widespread and large-scale devastation, and that was due to the tsunami it generated. The tsunami, rather than strong shaking, caused nearly all of the damage and loss of life in the Tohoku-Oki earthquake as well. There are several factors that mitigate the damaging effects of these earthquakes somewhat. One is the depth of faulting, which is several 10 s of kilometers. Another is that the strongest shaking in an earthquake is usually in the direction of rupture, which for typical subduction geometries means that the strongest shaking can be expected to occur updip and offshore.

There are emerging evidences that the area of strongest high-frequency wave propagation in these extremely large subduction zone earthquakes is the downdip end and that when rupture approaches the surface under the deep ocean, it may slip massively but without generating strong high-frequency seismic waves (Lay et al., 2012). There are several mechanisms that may contribute to this behavior including the interaction of a shallow-dipping fault with the free surface (Kozdon and Dunham, 2013), possibly enhanced by relatively compliant hanging wall material (Ma and Beroza, 2008), strongly velocity-weakening fault material (Shibazaki et al., 2011), and pore fluid pressurization (Mitsui et al., 2012).

A particular kind of earthquake occurs in a region known as the 'outer rise.' The outer rise is an area of relatively high topography, seaward of the trench, caused by flexural stresses from the downward bending of the subducting plate. Outer-rise earthquakes are thought to occur in response to these localized stresses in the subducting plate (Chapple and Forsyth, 1979). Despite the fact that convergent plate boundaries accommodate compressional deformation, both reverse and extensional normal faulting earthquakes are known to occur in the outer rise. These outer-rise earthquakes can be quite large. The damage due to shaking from them is usually less severe than for earthquakes on the subduction interface because they are more distant from land; however, outer-rise earthquakes can give rise to large and devastating tsunamis, such as in the case of the 1933 M_w 8.4 Sanriku event.

In some subduction zones, such as Japan, there is considerable deformation of the overriding plate as well. In the case of Japan, the upper-plate deformation is largely strike slip and is widely distributed. In cases where the plate convergence direction is highly oblique to the subduction zone, the slip is usually partitioned into trench-normal reverse slip on the subduction megathrust and trench-parallel strike slip in the overriding plate (Fitch, 1972; Yu et al., 1993). Several explanations have been offered for this partitioning. Savage (1983) suggested that the downdip edge of the coupled zone would lead to localized horizontal shear arising from the strike-slip component of the relative plate motion at that locale. Beck (1983) suggested that shear strain in the overriding plate was controlled by a weak zone corresponding to the volcanic arc. Another explanation is that slip partitioning would occur on distributed fault systems if each fault will have a smaller total area than one large fault. Thus, the total resisting force, which is the product of the frictional strength and the fault area, will

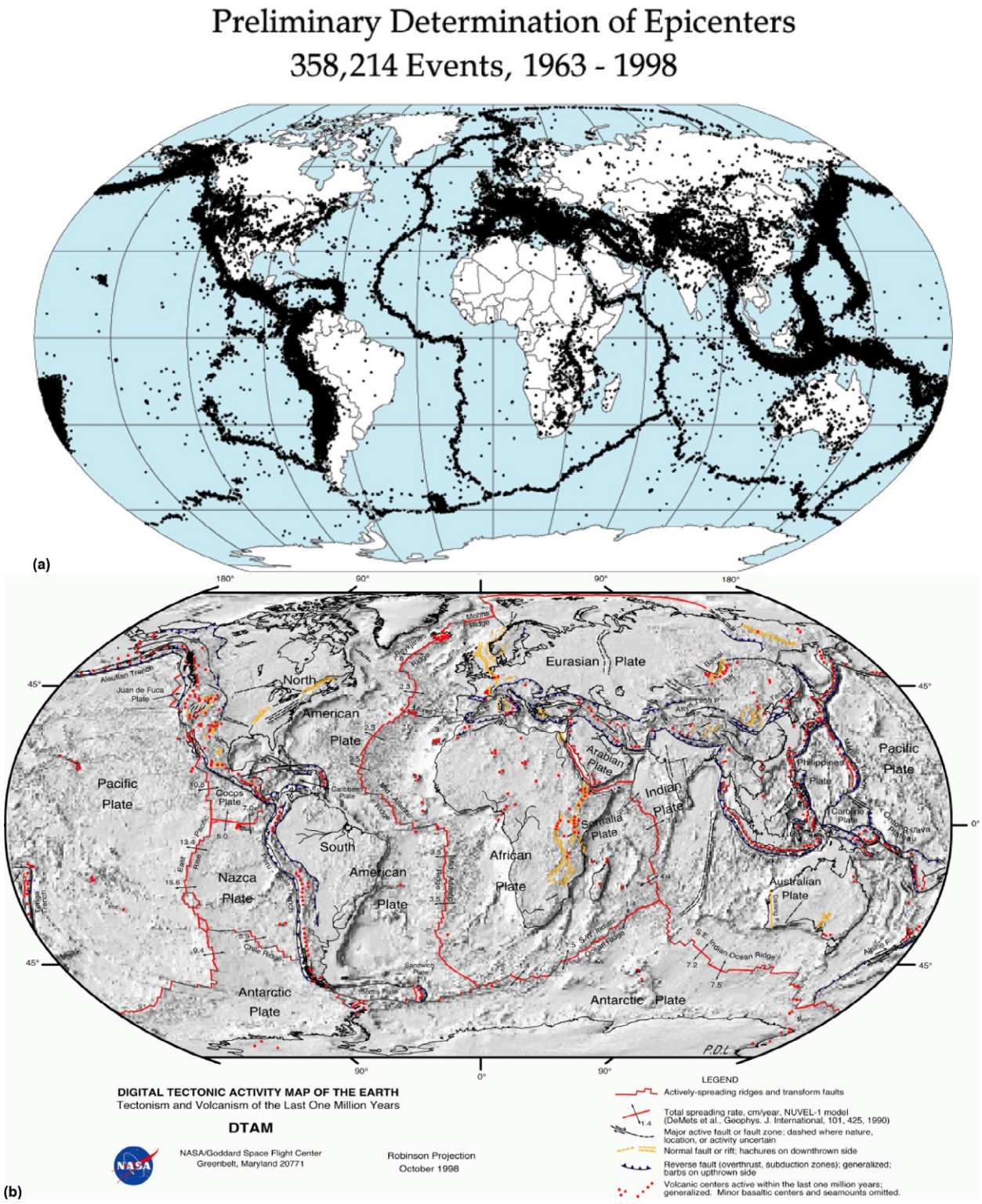


Figure 8 (a) Seismicity of the Earth for a 25-year time interval. Each black dot represents an earthquake and shows how earthquake activity limns global tectonic plate boundaries (reproduced from http://denali.gsfc.nasa.gov/dtam/seismic/). (b) Topography/bathymetry and the major tectonic plate boundaries of the Earth for a 25-year time interval. Note the close correspondence between plate boundaries and earthquake activity in **Figure 1(a)** (reproduced from NASA via Wikipedia Commons).

be lower, and hence, less force will be required to cause fault slip (McCaffrey, 1992). In subduction zones with oblique convergence, the strike-slip component of motion can be accommodated with less total frictional resistance if it occurs on a vertical strike-slip fault. Of course, this comes at the expense of having slip in more earthquakes. The trade-off between the two factors has been explored by Michael (1990), who determined simple geometries for which partitioning was energetically favorable.

Intermediate and deep-focus earthquakes, usually defined as earthquakes at depths of 70–300 km and 300–680 km, respectively, occur exclusively at convergent plate boundaries within subducting lithosphere. **Figure 9** shows the global depth distribution of intermediate and deep-focus earthquakes. In some cases, the subduction zone that gave rise to the lithosphere at depth may no longer be active at the surface, such as for the Spanish deep-focus earthquake of 1954 (Chung and Kanamori, 1976). The mechanism of intermediate and deep-focus earthquakes is a topic of continuing debate as Earth materials at such great pressures and temperatures are expected to deform plastically; however, several mechanisms have been put forward to explain their occurrence.

Deep earthquakes are most numerous in places such as Tonga, where cold material is rapidly subducted to great depth. The Tonga-Kermadec region accounts for nearly 2/3 of the seismicity of > 300 km depth. The largest known deep-focus earthquake is the M_w 8.3 Bolivia event of 1994 (Wu et al., 1995). A thorough investigation of deep earthquakes can be found in **Chapter 4.13**. **Figure 10** shows the depth distribution of earthquakes in the Tonga subduction zone.

The spatial distribution of intermediate-depth earthquakes is less well understood than that of deeper earthquakes, though plate bending and subduction zone coupling appear to play important roles (Astiz et al., 1988). The largest known intermediate-depth earthquake is the magnitude 8.1 Ryukyu Islands event of 1911. There is a great deal of variation in the behavior of intermediate-depth earthquakes. Intermediate-depth earthquakes reach magnitudes of at least M_w 7.8 (Astiz et al., 1988). For both intermediate and deep-focus earthquakes, we lack the kinematic plate tectonic boundary conditions that enable us to constrain the long-term probabilities of interplate earthquakes; hence, long-term forecasting of such earthquakes is by necessity based solely on statistical empiricism.

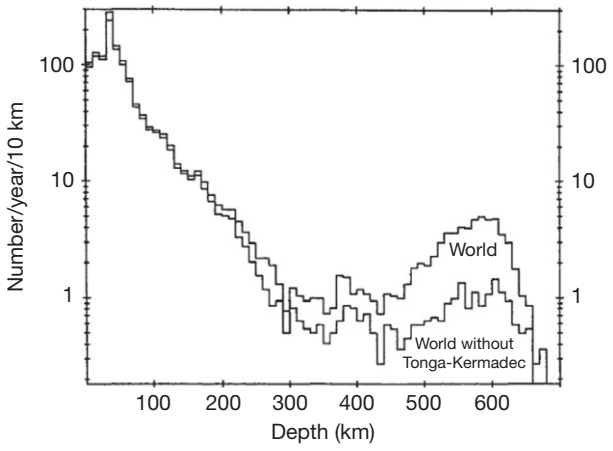


Figure 9 Depth distribution of earthquakes for the period 1/1964–2/1986 with mb five and larger. There is a steady decrease of seismicity up to ~300 km depth followed by a marked increase in seismicity below 500 km, with a maximum at ~600 km depth. The distribution is similar whether or not the prolific source of deep earthquakes in the Tonga–Kermadec region is included. Reproduced from Frohlich C (1989) The nature of deep focus earthquakes. *Annual Review of Earth and Planetary Science* 17: 227–254.

Some convergent plate boundaries involve continental crust on both plates. This occurs most notably along the Alpide belt, which includes the collision of Africa, Arabia, and India with southern Eurasia. Although in places, such as the eastern Mediterranean, this involves the subduction of small relict ocean basins, for the most part, continental collision zones entail the collision of relatively buoyant and weak continental crust that is not easily subducted. Deformation in these regions can extend many 100 s of kilometers from the main collision zone. In East Asia, the buoyancy of continental crust leads to the rise of the Himalayas and the Tibetan Plateau, as well as large strike-slip fault systems that allow buoyant crust to effectively flow away from the convergence zone. These zones give rise to some of the largest known strike-slip earthquakes, such as the great M_w 8.3 1905 Mongolia earthquake (Okal, 1992), and include major strike-slip fault zones, such as the North Anatolian Fault in Turkey.

As with earthquakes, much of the systematics of subduction zone volcanism was recognized before the advent of plate tectonics. Island arc volcanoes occur over subducting or recently subducted slabs. The melt that drives island arc volcanism is created when phase changes in the subducting slab give off water and other volatiles. When these come in contact with the surrounding mantle, the melting point is lowered, which in turn creates a buoyant melt that ascends and results in explosive volcanism. For detailed explanation of volcanic processes, see **Chapter 4.14**.

4.01.2.2.4 Intraplate earthquakes

Intraplate earthquakes occur far from plate boundaries and thus are not explained by the theory of plate tectonics. We distinguish here between earthquakes that occur within broad zones of plate boundary deformation, such as those in the Basin and Range province of western North America, and truly intraplate earthquakes, such as those in the New Madrid seismic zone of central North America. Earthquakes that fall into a sort of gray area between plate boundary earthquakes and true intraplate earthquakes include earthquakes in central and southern India and earthquakes in the East African Rift system. Locales of true intraplate earthquakes include Central Australia, Europe's Rhine Graben, Central Brazil, and large parts of oceanic crust under the Indian and Pacific Oceans. The M_w 8.6 11 April 2012 earthquake off the coast of Sumatra was a borderline intraplate event. It occurred under the Indian Ocean, offshore from the 2004 Sumatra earthquake, but the region in which it occurred had previously been identified as a potentially incipient plate boundary between the Australian and Capricorn plates (Royer and Gordon, 1997).

Why do intraplate earthquakes occur? About half of all intraplate earthquakes are observed to occur in failed

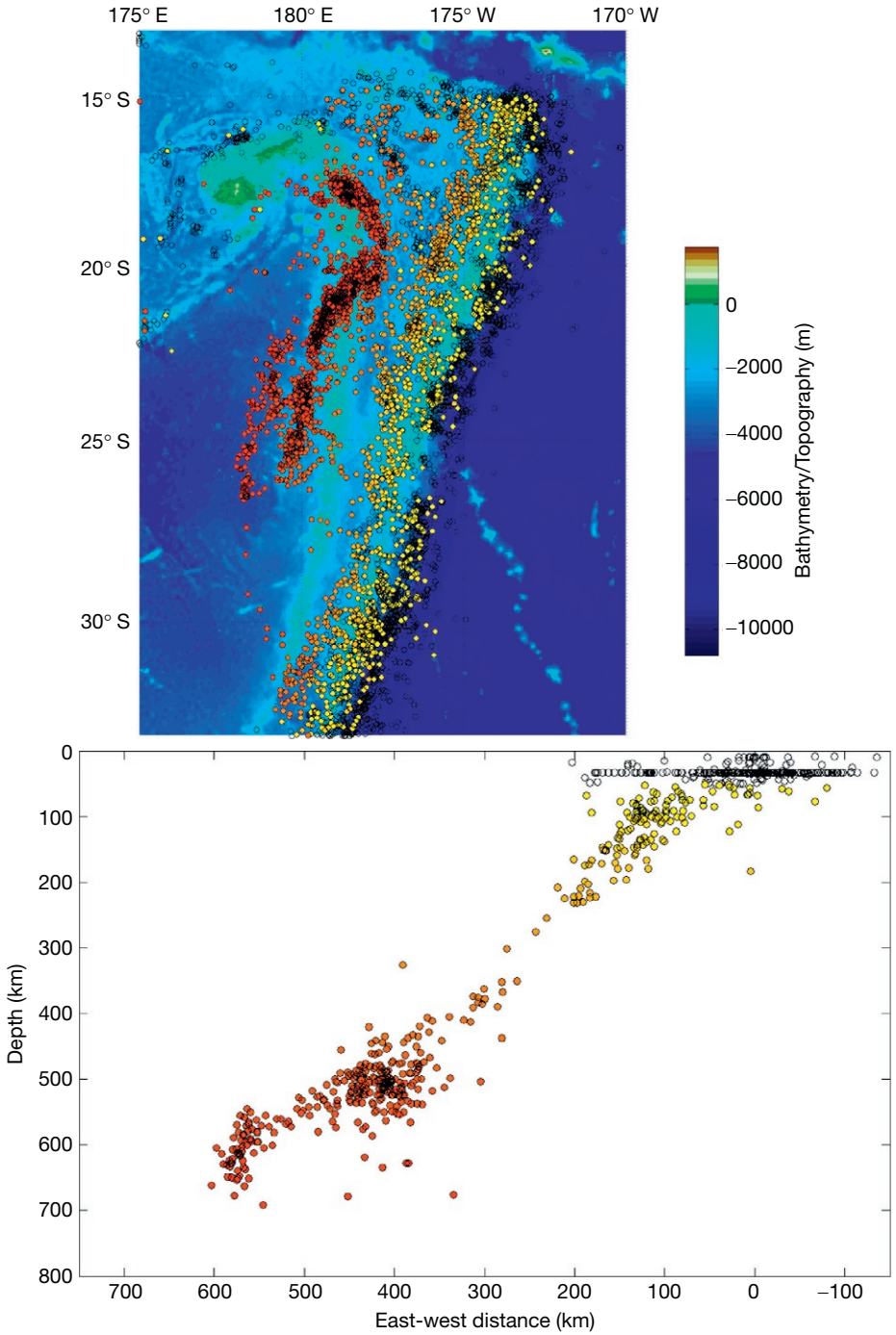


Figure 10 Top panel shows $M > 5$ seismicity (from CISN catalog for 1964–2006) in the Tonga region color-coded by depth of hypocenter. Tonga Trench is dark blue lineation in bathymetry. Open symbols represent earthquakes < 50 km depth, yellow 50–100 km depth, and increasing red tone denotes increasing depth up to nearly 700 km. Horizontal lineation at 15 km is default depth used when depth is indeterminate. Lower panel shows E–W cross section at (24–27°S latitude). Rapid subduction of old lithosphere results in the world's most prolific source of intermediate and deep-focus earthquakes.

continental rifts (Johnston and Kanter, 1990). These are likely to be weaker than the unrifted continental crust around them, so it is reasonable to conclude that such intraplate earthquakes exploit this relative weakness. This does not appear to provide a universal explanation, however, as intraplate earthquakes in places such as Australia and oceanic plates do not occur on such structures. Intraplate earthquakes respond to the ambient stress field, which supports the possibility that they are occurring due to relative lack of strength, rather than excess of stress. Other explanations for the driving mechanism behind intraplate earthquakes include stresses induced by the emplacement or sinking of plutons (Kane, 1977), postglacial rebound

(Arvidsson, 1996), and stress concentration due to the presence of a weak zone in otherwise strong crust (Campbell, 1978; Sbar and Sykes, 1973).

Regardless of their mechanism, it is clear that intraplate earthquakes do not account for a very large fraction of lithospheric deformation. If they did, then global plate motion models would fail closure tests. Strain rates due to intraplate deformation have an upper bound of 4×10^{-10} year^{-1} (Gordon, 1998). Johnston and Kanter (1990) found that the cumulative seismic moment of intraplate earthquakes was about 0.5% of that of the Earth's total. Despite this small percentage, intraplate earthquakes do pose a significant seismic hazard and are occasionally quite large, such as in the 1886 Charleston earthquake, with an estimated moment magnitude of 7.6, or the New Madrid earthquake sequence of 1811–12, which by some estimates included three earthquakes of moment magnitude approximately equal to 8. More recent estimates of the magnitude of the large earthquakes in the New Madrid sequence are closer to 7.0 (Hough and Page, 2011).

4.01.2.2.5 Hot spot volcanism

As with intraplate earthquakes, there are important types of volcanoes for which plate tectonics does not have a ready explanation. These are known as hot spot volcanoes – areas of profuse volcanism that persist for 10 s of millions of years. Examples of hot spot volcanoes include some of the most famous and well-studied volcanic systems on Earth, such as Hawaii, Kerguelen, and Iceland. Yellowstone is a classic example of a continental hot spot volcano. Hot spot volcanism in the Earth's oceans tends to be characterized by effusive basaltic eruptions, whereas the interaction of basaltic magma with continental crust can lead to large-scale explosive rhyolitic volcanism. The ultimate magmatic source of hot spot volcanism remains a matter of some debate (Montelli et al., 2004), with some evidence pointing to a source in the lowermost mantle, while other evidence favors a shallower origin.

4.01.2.3 The Largest and Deadliest Earthquakes

The 12 largest instrumentally recorded earthquakes (updated from Kanamori, 1977) are listed in **Table 2**. All but one of them are plate boundary earthquakes in subduction zones. The largest instrumentally recorded earthquake was the 1960 Chile earthquake, which had a moment magnitude of 9.5 (instrumentation adequate for calculating reliable magnitudes has only been available since about 1900). The 1964 Good Friday earthquake in Alaska had a surface wave magnitude of 8.4 and a moment magnitude of 9.2. Put another way, the Alaska earthquake was 20 times larger than would be estimated from the surface wave magnitude. The 1989 Loma Prieta earthquake had a moment magnitude of 6.9. For comparison, the 1960 Chile earthquake involved nearly 7000 times the seismic moment of the Loma Prieta earthquake. The 1964 Alaska earthquake had nearly 3000 times the potency of Loma Prieta. The 1906 San Francisco earthquake had a moment magnitude of ~7.9 (this is more uncertain than the other estimates due to the sparse instrumentation at the time); hence, the seismic moment for that earthquake was about 30 times higher than for the Loma Prieta earthquake. There is no evidence for events larger than the 1960 earthquake in the geologic record, and the extreme length, width, and slip of the 1960 Chile earthquake are impressive; however, given that the instrumental era in seismology spans only slightly more than 100 years, yet larger earthquakes have to be considered possible.

Table 2 The 12 largest earthquakes

Year	*Location*	M_w
1960	Chile	9.5
1964	Alaska	9.2
2004	Sumatra	9.2
2011	Japan	9.0
1952	Kamchatka	9.0
2010	Chile	8.8
1906	Ecuador	8.8
1965	Aleutians	8.7
2005	Indonesia	8.6
2012	Off-Sumatra	8.6
1957	Aleutians	8.6
1950	Assam	8.6

Half of the 12 largest recorded earthquakes have occurred in the past 10 years, which suggests a nonrandom clustering of great events. Several tests of significance, however, reveal that this flurry of activity does not represent a significant deviation from Poisson (random in time) behavior (Michael, 2011; Shearer and Stark, 2012).

Magnitude scales quantify the physical size and geologic impact of earthquakes, but the impact of earthquakes on human affairs depends strongly on where, when, and how they occur. **Table 2** lists the ten deadliest earthquakes. In the case of the 2004 Sumatra earthquake, nearly all of the fatalities were caused by the tsunami that the earthquake generated. In the case of the 1923 Kanto earthquake, the firestorm that followed played a major role (**Figure 11**). In the 1556 Shaanxi and 1920 Gansu earthquakes, the widespread collapse of caves used as dwellings and carved into poorly consolidate loess magnified the losses tremendously. Five of the 12 deadliest earthquakes in history occurred in China, which is attributable to a long history of dense population in an earthquake-prone region (**Table 3**).

It is difficult to assess how reliable these numbers are. Certainly, for earthquakes that occurred 100 s of years ago, the reliability can be questioned. Even in the modern age, the devastation and confusion following such massive catastrophes, and the need to focus on recovery rather than counting fatalities in their immediate aftermath, mean that the true death tolls for these events will never be known accurately. For example, given the widespread and extreme devastation caused by the 2004 Indian Ocean earthquake and tsunami, the number of people lost in that catastrophe will never be known. Moreover, in some cases, political sensitivities may have influenced the reporting of casualties, contributing to the uncertainty. Unofficial estimates of the death toll of the 1976 Tangshan earthquake (**Figure 12**) range as high as 655 000. Following the 1999 Izmit, Turkey, earthquake (not listed), authorities stopped counting when the death toll reached ~18 000. Using several lines of evidence, Marza (2004) suggested a lower bound of the death toll in this earthquake

Figure 11 Print depicting the firestorm that followed the 1923 Kanto, Japan, earthquake. The earthquake and firestorm that followed devastated the Tokyo–Yokohama metropolitan area and are estimated to have killed 143 000 people. Reproduced from http://www.rekihaku.ac.jp/e-rekihaku/109/pic19.jpg.

Table 3 The 12 deadliest earthquakes

Year	*Location*	*Fatalities*
1556	Shaanxi, China	830 000
2004	Sumatra, Indonesia	283 106
1976	Tangshan, China	255 000
1138	Aleppo, Syria	230 000
2010	Port-au-Prince, Haiti	228 000
856	Damghan, Iran	200 000
1927	Tsinghai, China	200 000
1920	Gansu, China	200 000
893	Ardabil, Iran	150 000
1923	Kanto, Japan	143 000
1948	Turkmenistan, USSR	110 000
2008	Wenchuan, China	87 000

as 45 000. There is considerable uncertainty about the loss of life in the 2010 Haiti earthquake as well. Whatever their precise toll, one thing is clear: earthquakes can be devastating. At their worst, they rank among the most extreme, sudden catastrophes humankind has to cope with.

Due to the rapid growth of the Earth's population, the concentration of that growth on the earthquake-prone areas of the Pacific Rim, and the urbanization accompanying that growth, vulnerability to earthquakes is growing. This is true despite steady and considerable progress in earthquake-

Figure 12 Photo showing nearly complete devastation in the Lunan District of the city of Tangshan from the 1976 Tangshan, China, earthquake (Huixian et al., 2002). Tangshan was a major industrial center with a population of one million before the earthquake, which occurred directly under the city at 3:42 a.m. About 93% of residential buildings and 78% of commercial buildings in the city were destroyed (Yong et al., 1988). The officially reported death toll is 255 000, but estimates are as high as 655 000.

resistant engineering. Bilham (1995) suggested that an earthquake within the next 30 years with over a million fatalities is not out of the question.

Loss of life represents the most important vulnerability to earthquakes, but it is not the only one. Economic losses in earthquakes can be grave as well. For example, by some estimates (Risk Management Solutions Inc, 1995), a repeat of the 1923 Kanto, Japan, earthquake would likely result in direct economic losses of over 2 trillion dollars (44–70% of Japan's GDP) – a financial event that would be without historical precedent and that would have far-reaching global consequences.

4.01.2.4 Historic and Prehistoric Earthquakes

For the purposes of discussion, we define 'historic' to denote an earthquake for which there are written accounts, but that occurred before the dawn of instrumental seismology in the last decade of the nineteenth century, so that there is no instrumental record of ground motion. We use the term 'prehistoric' to denote an earthquake that occurred so long ago that there is no written record of it.

Some earthquakes have had a significant impact on the course of history. Crushed skeletons in ruins of the ancient world provide compelling evidence that destructive earthquakes occurred in the cities of Troy, Megiddo, and Jericho. They may be the inspiration for the Old Testament story about the fall of the walls of Jericho. Megiddo (Armageddon) is thought to have been the inspiration for the biblical story of the apocalypse. The 1755 Lisbon earthquake made a deep impression on philosophers in eighteenth-century Europe. The earthquake occurred on a Catholic holiday, destroyed many churches and much of the city through a combination of strong shaking, fire, and tsunami. The Lisbon earthquake (Figure 13) coincides in time with, and may have helped bring about, the decline of Portugal as a major world power. Some have argued that the destruction and chaos brought about by the devastating 1923 Kanto earthquake led to fervent Japanese nationalism that helped fuel World War II. The expropriation of international relief funds for recovery from the 1972 Managua, Nicaragua, earthquake is thought to have helped bring about the communist takeover of that country shortly afterward. The devastation wrought by the 1976 Tangshan earthquake was followed within months by the death of Mao Tse-tung and the end of the Cultural Revolution. For a more detailed discussion on ancient earthquakes and the role of earthquakes in history, see Chapter 4.22.

Prehistoric earthquakes are in some cases studied using archeology. It is far more common, however, to study them using geology. This field of earthquake science has come to be known as paleoseismology. The field of paleoseismology allows geologists to address questions, such as whether or not a fault is active and what the nature of earthquake recurrence is, that are difficult or even impossible to address by other means. For a detailed discussion of paleoseismology, see Chapter 4.21.

4.01.2.5 Earthquake Size Distribution

Small earthquakes are much more common than large earthquakes. This statement is usually quantified in seismology using the frequency–magnitude relation – that is, an equation that describes the relative frequency of large versus small earthquakes. The frequency–magnitude relation is widely observed to follow Gutenberg–Richter statistics, named for the same seismologists who developed the magnitude scale. If $N(M)$ is the number of earthquakes of magnitude M or larger, then frequency–magnitude relation can be expressed as

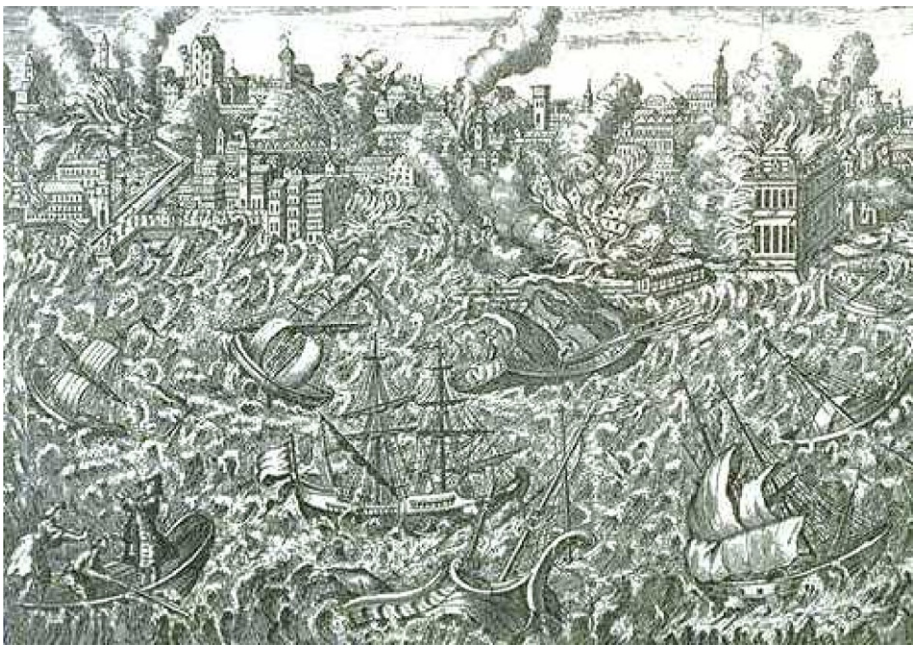


Figure 13 Copper engraving depicting the 1755 Lisbon, Portugal, earthquake of 1 November 1755 with fire ravaging the city and a tsunami destroying ships and the waterfront.

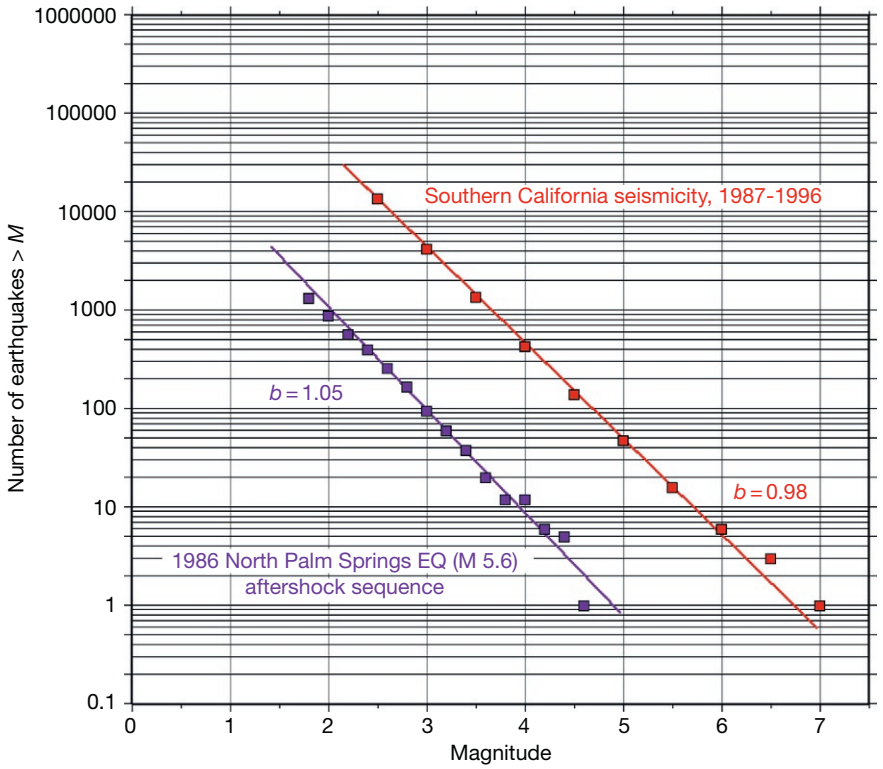


Figure 14 Fit of the cumulative form of the Gutenberg–Richter distribution for two earthquake populations: the 1986 North Palm Springs aftershock sequence (mainshock removed) and a decade of seismicity in all of Southern California. In both cases, the *b*-value is ~1. Reproduced from http://www.data.scec.org/Module/Pics/s2inset1.gif.

$$\log N(M) = a - bM$$

(Gutenberg and Richter, 1956). Two examples of earthquake populations that follow such a frequency–magnitude relation are shown in **Figure 14**. This linearity in semilog space describes a power law, with the *a*-value (intercept) corresponding to the number of magnitude 0 or larger earthquakes in the population considered and the *b*-value (slope) describing the relative number of large versus small earthquakes. While, the *a*-value is dependent on the population size and hence has no obvious physical significance, the *b*-value describes the relative frequency of large and small events regardless of the population size. A *b*-value of 1 means that for each unit magnitude increase, there is a factor of 10 reduction in the number of earthquakes.

For many earthquake populations, including the global seismicity dataset, a *b*-value of ~1 describes the relative numbers of large versus small earthquakes. For individual faults, however, the frequency–magnitude relation may fail to describe earthquake populations adequately (Wesnousky, 1994), and to the extent that they do, there is evidence that the *b*-value for an individual fault may decrease systematically with the maturity of a fault system (Stirling et al., 1996). For volcanic systems, *b*-values are typically much higher than for tectonic faults, sometimes reaching values as large as three (McNutt, 2005), indicating that the seismicity associated with active magmatic systems can be extraordinarily rich in very small earthquakes.

By combining the definition of the moment magnitude with the frequency–magnitude relation, it is straightforward to demonstrate that the largest earthquakes on a fault system dominate the slip budget. Combining the frequency–magnitude relation with the definition of the moment magnitude yields the following proportionality between the cumulative seismic moments as a function of earthquake magnitude, *M*:

$$\sum M_0(M_W) \sim 10^{(1.5-b)M_W}$$

From this, it is apparent that if the *b*-value is <1.5, the largest earthquakes will dominate the seismic moment budget. For the globally representative value of $b=1$, the seismic moment contribution increases by a factor of $10^{0.5}$, or ~3.2, with each increase in magnitude. The relationship between magnitude and seismic energy has the same dependence (Gutenberg and Richter, 1942). Thus, as was recognized by Richter in his original paper on earthquake magnitude, nearly all the energy is released by the largest earthquakes. A direct consequence of this is that little earthquakes do not relieve appreciable strain energy in large fault systems and cannot act as a 'safety valve' for large earthquakes.

4.01.2.6 Earthquake Location

An earthquake's hypocenter is defined as the temporal and spatial coordinates where seismic waves are first generated. Earthquake location has long been routine, but there continues to be important progress in developing more precise earthquake locations. The fundamental observations used to estimate earthquake locations are seismic wave arrival times. Over

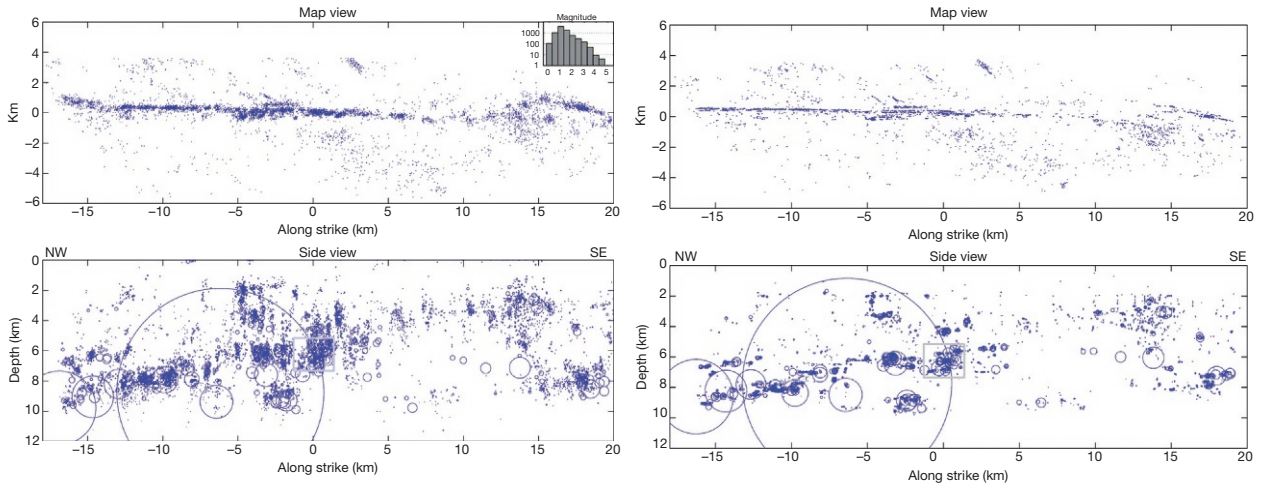


Figure 15 Seismicity on the Calaveras Fault before (left) and after (right) double-difference earthquake relocation using cross correlation-based arrival time measurements (Schaff et al., 2002). Both map and side views are shown for the two cases, which show nearly 8000 earthquakes. Earthquakes are shown as circular sources assuming constant stress drop. The improved locations are evident in the narrow width of fault structures at depth and in the large number of repeating earthquakes found on this fault. Such improved locations allow testing of diverse hypotheses regarding earthquake fault mechanics.

the last decade, precise earthquake location methods (e.g., Waldhauser and Ellsworth, 2000) coupled with the use of waveform cross correlation on large waveform databases to reduce measurement error (e.g., Schaff et al., 2004) have revolutionized our ability to resolve the fine structure of fault zones (**Figures 15** and **16**).

By using waveform cross correlation on large numbers of similar waveforms, seismologists have been able to reduce the uncertainty in arrival time measurements. For similar earthquakes, it is straightforward to reduce the measurement error of arrival times to several milliseconds for regional seismic networks. Since the measurement error maps directly into location uncertainty, the result is greatly improved earthquake locations. Arrival time measurement using cross correlation is not a new idea in seismology (Poupinet and Ellsworth, 1984), but only recently with the widespread availability of massive waveform datasets and computers capable of analyzing them has this approach reached its full potential.

The second advance that has led to improved locations is the expansion of the use of joint hypocentral location methods, particularly ones that are designed to minimize the effects of Earth structure on the solutions. These methods focus on precision (i.e., relative earthquake locations) rather than accuracy (absolute earthquake locations). These methods solve for the position of a large population of earthquakes simultaneously and when used with differential arrival time measurements allow seismologists to reduce the effects of unmodeled velocity structure on arrival times. Through such techniques, which are analogous to adaptive optics in astronomy, it is possible to reduce earthquake location errors dramatically. For the example shown in **Figure 15**, Schaff et al. (2002) documented a reduction in hypocentral uncertainty by an order of magnitude in their relocations. As with large-scale waveform cross correlation, enhanced computational capability has been an essential element in large joint hypocentral determinations.

An intriguing finding based on highly precise earthquake locations is the discovery of slip-parallel 'streaks' of seismicity (Rubin et al., 1999; **Figure 16**). These streaks have been observed in diverse environments, including the southeast rift zone of Kilauea volcano, Hawaii, as well as the San Andreas, Calaveras, and Hayward Faults in California. Waldhauser et al. (2004) examined several such streaks along the Parkfield segment of the San Andreas Fault and concluded that, in one case, they may mark the boundary between locked and slipping parts of the fault, while in another case, they are more easily explained as a result of a discontinuity in geologic structure.

4.01.3 The Earthquake Source

Most of the measures of earthquake size and the determination of earthquake locations that we have discussed in the previous section, with the exception of the seismic moment, do not require a quantitative theory of the earthquake source. In this section, we outline some of the applications of earthquake source theory to seismology. **Chapter 4.02** provides a much more detailed and complete treatment.

4.01.3.1 Point-Source Parameters

Quantitative studies of the earthquake source usually treat the Earth as an elastic continuum, with the exception of the fault itself, which is assumed to be a discontinuity having infinitesimal width. These are both approximations that are not uniformly valid and must be kept in mind when interpreting source characteristics. The source of nearly all earthquakes is shear slip on faults. To first order, faults represent planar discontinuities within the Earth and the shear slip occurring on them is parallel to that fault surface. Equation [45] of **Chapter 4.02** expresses the displacement seismogram at a point, x, as a function of time, t, as an integral over the fault surface and time of the fault slip history, $\Delta u_i(x_0,\tau)$:

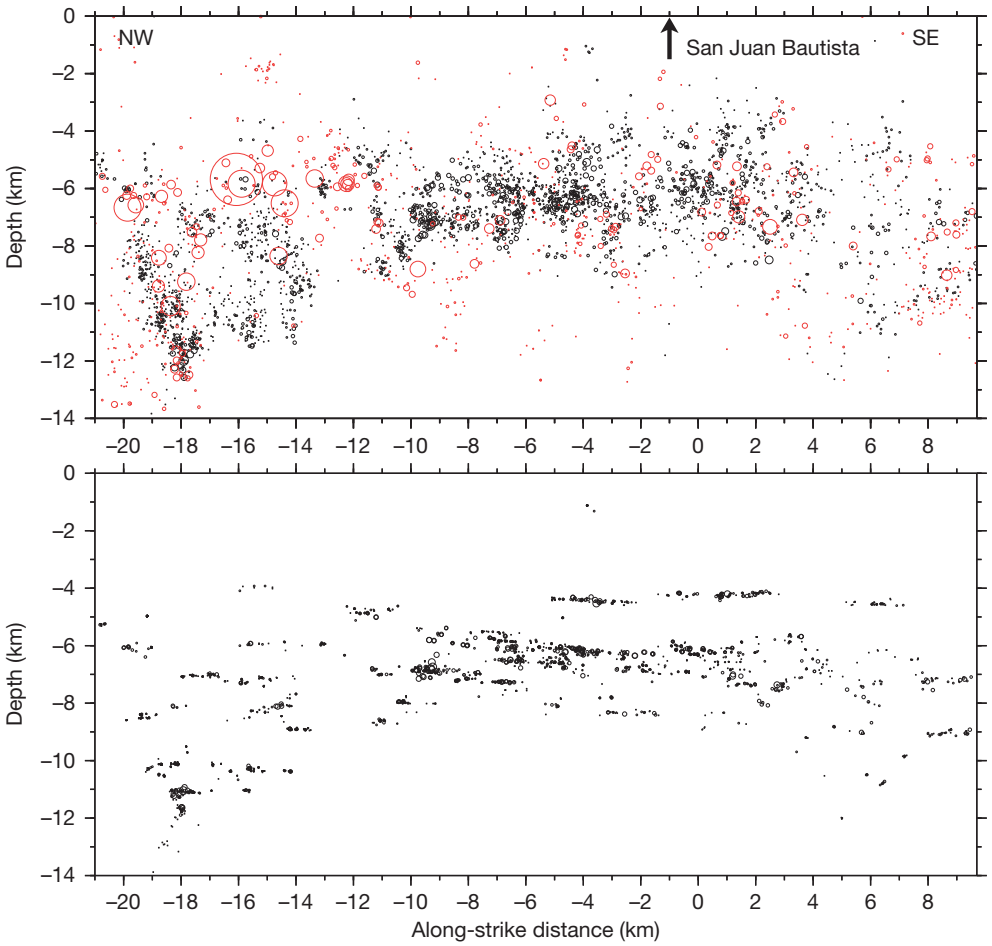


Figure 16 Network locations of microearthquakes on a vertical cross section along the San Andreas Fault near San Juan Bautista (Rubin et al., 1999) are shown on the upper panel. Red circles in upper panel represent larger events that were not relocated because waveforms were clipped. Circles denote estimated slip area for larger events. Lower panel shows same events after relocation. The horizontal seismicity streaks, which were previously obscured by location errors, emerge in the relocated catalog.

$$u_i(x,t) = \int_0^t \int_{s_{x0}} \mu(x_0)\Delta u_j(x_0,\tau)G_{ij,k}(x,t|x_0,\tau)n_k(x_0)\mathrm{d}^2x_0\mathrm{d}\tau$$

The other terms in the integral – the shear modulus, μ; Green's function gradient, $G_{ij,k}$; and the fault-plane normal vector, n_k – together represent the traction across the fault surface due to an impulsive point source acting at the receiver.

If the wavelengths considered are much larger than the rupture dimension, and if we record seismic waves sufficiently far away from the source, then we can make the point-source approximation. In this case, variations in Green's function, G, with position on the fault, x_0, are negligible and the integration mentioned earlier becomes trivial. Other approximations include a high-frequency approximation, wherein only the far-field $(1/r)$ terms in Green's function make important contributions to the seismogram. This approximation is valid once the receiver is several wavelengths from the source $(\omega r/\beta \gg 1)$ and allows the integral mentioned earlier to be simplified since only terms that can be calculated using ray theory need be retained. An intermediate assumption, analogous to the Fraunhofer approximation in optics, is that Green's function varies over the fault plane, but only in phase. This leads to a simplification of the integral mentioned previously into one in which the G is outside the integral and results in an expression for the far-field pulse shape, which represents the dependence of the far-field wave amplitude with time. The development of the theory and approximations outlined earlier are covered in greater detail in Chapter 4.02.

4.01.3.2 Sense of Faulting from First Motions

The polarity of the first-arriving P-waves can be used to determine a focal mechanism, which is a representation that plots the polarity of waves as a function of takeoff angle from the source. For a shear-slip source, the P-waves will have four lobes, two of which have compressional first-arriving P-waves and two of which have dilatational first-arriving P-waves (Figure 17). At 45°, from these lobes are two planes for which the P-wave amplitude is nodal. One of these planes is the true fault plane of the earthquake, while the other is known as the 'auxiliary' plane. From the polarity of the P-waves, it is impossible to distinguish these two planes from one another. Indeed, in the point-source approximation, the entire seismogram, not just the first motion polarity, is identical for slip on either one of these planes. This leads to a fundamental fault-

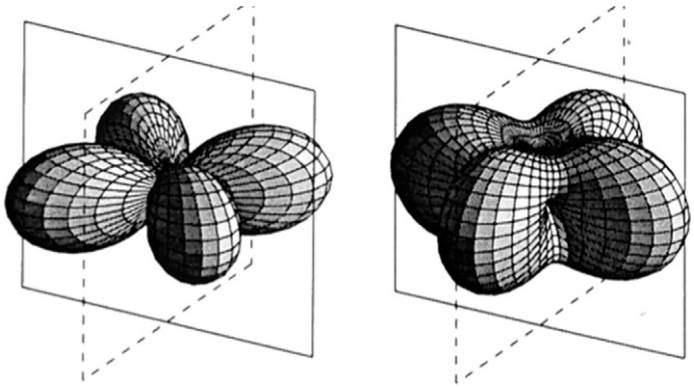


Figure 17 Radiation patterns for a strike-slip fault showing the amplitude of P-waves on the left and the amplitude of S-waves on the right. The four-lobed P-wave radiation pattern has two compressional and two dilatational lobes. The fault plane and auxiliary plane, for which P-wave amplitude is nodal, separate the four lobes. The lobes of the radiation pattern are separated by quadrants in the usual depiction of the focal mechanism. Reproduced from Julian BR, Miller AD, and Foulger GR (1998) Non-double-couple earthquakes 1. Theory. *Reviews of Geophysics* 36: 525–549.

plane/nodal-plane ambiguity. Without independent information, such as geologic information on the fault orientation, alignment of earthquake hypocenters, or finite-source effects in the waveforms (i.e., a breakdown of the point-source assumption), it is impossible to tell the two planes apart. The sense of faulting as determined using the P-wave radiation pattern using earthquake focal mechanism solutions provided a key test of the theory of plate tectonics in which the sense of motion on oceanic transform faults was found to be consistent with the plate tectonic prediction (Sykes, 1967).

4.01.3.3 Moment Tensor Representation

A more general representation of a point source of seismic waves is provided by the seismic moment tensor. The moment tensor can be used to represent the seismic source as a pair of force couples – the so-called double-couple solution, which is the point-source equivalent force distribution for a dislocation source. The determination of seismic moment tensors for earthquakes is now routine such that there exist catalogs of moment tensor solutions at both the regional scale and the global scale. Perhaps, the best known of these is the Harvard Centroid Moment Tensor (CMT) catalog (Dziewonski et al., 1981) and its successor the Global CMT Project, which contains moment tensor solutions for over 20 000 earthquakes worldwide (**Figure 18**). Similar catalogs of earthquake mechanisms, whether based on first motion focal mechanisms or waveform-based moment tensor inversions, are available for thousands of locally recorded earthquakes where local seismic networks provide the data to make such work possible.

The strength of the moment tensor representation is that it can be used to describe sources other than shear slip on faults. Sources with isotropic (volumetric), or deviatoric, but not planar shear slip, have been studied using seismic waves, most often in volcanic environments. Single-force solutions have also been applied to describe the sources of seismic waves for large landslides, such as the one that initiated the 1980 eruption of Mt. St. Helens (Kanamori and Given, 1982). For more details on moment tensors and their application to the study of earthquakes, see

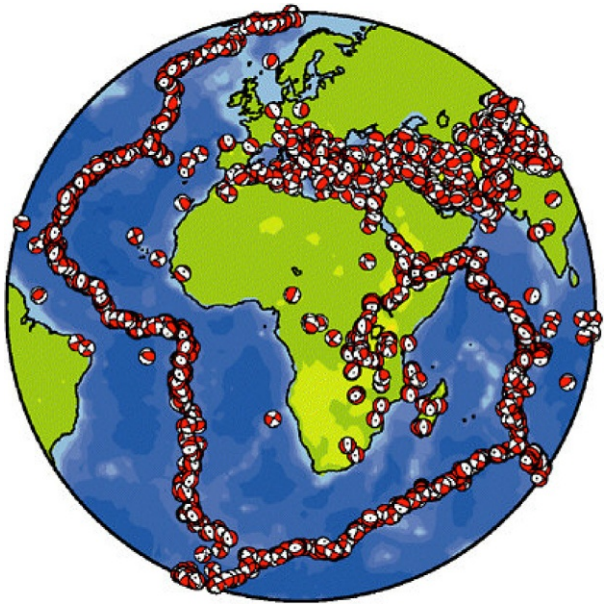


Figure 18 First motion focal mechanism representation of CMT solutions from the Harvard catalog for earthquakes around the continent of Africa for the 25-year period (1976–2000). Reproduced from National Research Council of the National Academies (2003) Living on an Active Earth, perspectives on earthquake science. Committee on the Science of Earthquakes Board on Earth Sciences and Resources Division on Earth and Life Studies. Washington, D.C.: The National Academies Press.

Chapters 4.02 and **4.18**. For applications to volcano seismology, see **Chapter 4.15**.

4.01.3.4 Seismic Energy

A point-source parameter that is important, but for which estimates are far from routine, is the radiated seismic energy. The radiated energy is typically calculated by measuring the energy flux at a point, or points, where the wave field of an earthquake is recorded, and from that inferring the total energy flux in all directions from the earthquake source. The radiated energy is broadband, with most of the energy being radiated at frequencies higher than a frequency known as the 'corner' frequency (Singh and Ordaz, 1994).

Seismic sources can be characterized by an overall length scale. Waves of sufficiently low frequency will have corresponding wavelengths that are much larger than that source dimension. Waves of sufficiently high frequency, on the other hand, will have corresponding wavelengths that are much smaller than that of source dimension. In this latter case, the amplitude of seismic waves will suffer from interference effects and hence have smaller amplitudes in the far field. The crossover frequency, at which wavelengths are comparable to the seismic source dimension, is known as the corner frequency. The corner frequency can be used to infer the characteristic size of the seismic source.

Strictly speaking, because the corner frequency defines the frequency at which finite-source effects start to become important, the seismic energy is not truly a point-source parameter. As a scalar measure of earthquake strength, however, it is natural to consider it as such. Obtaining the necessary bandwidth with sufficient signal-to-noise ratio, and correcting for propagation effects over that wide bandwidth, is what makes the radiated energy so difficult to calculate reliably (Ide and Beroza, 2001). The study of seismic energy, both its measurement and its implications for the physics of earthquakes, is the subject of an entire AGU monograph (Abercrombie et al., 2006).

4.01.3.5 Extended-Source Models of Earthquakes

The success seismologists have had in cataloging basic point-source earthquake parameters, particularly location, origin time, moment tensor, and seismic moment contrasts with attempts to estimate higher-order source parameters. This is natural in that these extended-source measurements involve either spatial or temporal dimensions, that is, slip duration, directivity, stress drop, and moment-rate function, that can be difficult to resolve. Methods to recover these parameters remain an area of active research.

Studies of the earthquake source are sometimes divided into 'kinematic' and 'dynamic' studies. The adjective 'kinematic' is used in seismology to describe models of an earthquake that estimate source parameters without explicit consideration of the failure mechanism on the causative fault. This is not necessarily a negative property. With our incomplete knowledge of fault mechanics, it is likely that dynamic models, which seek to incorporate what we think we know about earthquake physics, may be incorrectly biased by those assumptions. The hope is that observations of earthquake kinematics will inform dynamic models of fault behavior and that dynamic models will motivate kinematic modeling that tests dynamic hypotheses. We start our discussion of kinematic source models with point-source models.

4.01.3.5.1 Kinematic source models

The moment tensor representation that is used to describe earthquakes in the point-source approximation can be extended to finite sources using an expansion in higher-order moments (Backus and Mulcahy, 1976). McGuire et al. (2002) applied this formalism to recover the degree-two moments of large earthquakes recorded at teleseismic distances. They found that in 80% of the earthquakes studied, rupture was unilateral (rupture initiated at one end of the rupture zone) rather than bilateral (rupture initiated near the center of the rupture zone). For the most part, seismologists have not applied the formalism of higher-order moments to the study of the earthquake source. This may be due to a desire to image the source directly as the spatial and temporal evolution of slip on a fault.

A great deal of work has gone into imaging earthquake rupture directly as slip on a fault (or faults) using seismic waves. For some earthquakes, important additional constraints have been provided by geodetic measurements or field measurements of surface rupture. **Chapter 4.09** details approaches taken for imaging earthquake rupture.

As shown by Ide et al. (2005), the nonuniqueness of the fault slip inverse problem is extensive. When there are insufficient near-source data, such as for the 1999 Izmit, Turkey, earthquake, slip models developed by different investigators bear little resemblance to one another. In contrast, well-recorded earthquakes like the 1999 Chi-Chi, Taiwan, earthquake will yield much more robust results. For all but extremely large events (Ishii et al., 2005), extended-source images are difficult to recover without near-source data. As a result, most of the information available on source finiteness derives from well-instrumented areas like Japan and California.

A large earthquake may rupture multiple-fault segments. Modeling of strong ground motion recordings of large earthquakes suggests that fault rupture typically propagates at about 80% of the shear-wave velocity, although there is evidence that the rupture velocity can locally exceed the shear-wave velocity (Bouchon and Vallée, 2003; Bouchon et al., 2001; Dunham and Archuleta, 2004). The slip velocity across the fault is much more difficult to constrain but is thought to be on the order of meters per second in a large event (Heaton, 1990). In addition to constraints on the evolution of slip in earthquakes, it is clear from all studies that have sufficient resolution that slip is strongly variable with position on the fault (Mai and Beroza, 2002). **Figure 19** shows examples of variable slip models for five different earthquakes.

Following the 2004 Sumatra earthquake, Ishii et al. (2005) pioneered a new approach to source imaging. Their back-projection approach uses coherent signals recorded across an array to stack coherent energy over a relatively narrow frequency band. It has distinct advantages over slip inversion because it more readily allows for multiple-segment or temporally complex rupture such as observed in the 2012 off-Sumatra earthquake (Meng et al., 2012); however, it does not image a physical quantity (fault slip) as traditional kinematic source inversions do. More on back-projection and kinematic source inversion can be found in **Chapter 4.09**.

4.01.3.5.2 Dynamic source models

Dynamic source models differ from kinematic models in that they explicitly consider the stress associated with fault slip and attempt to model stress, slip, and the evolution of rupture using a failure criterion. In some cases, the evolution of the rupture is prescribed, and the relevant physical parameters are calculated. Quasidynamic rupture models (Quin, 1990) fall within this category. In others, the fault is allowed to evolve based on the initial conditions and a failure criterion for the fault surface. A dynamic model such as this is referred to as a 'spontaneous' model because once initiated, the rupture propagates spontaneously as long as the energy balance allows it to be self-sustaining (Kostrov and Das, 1988). A third class of

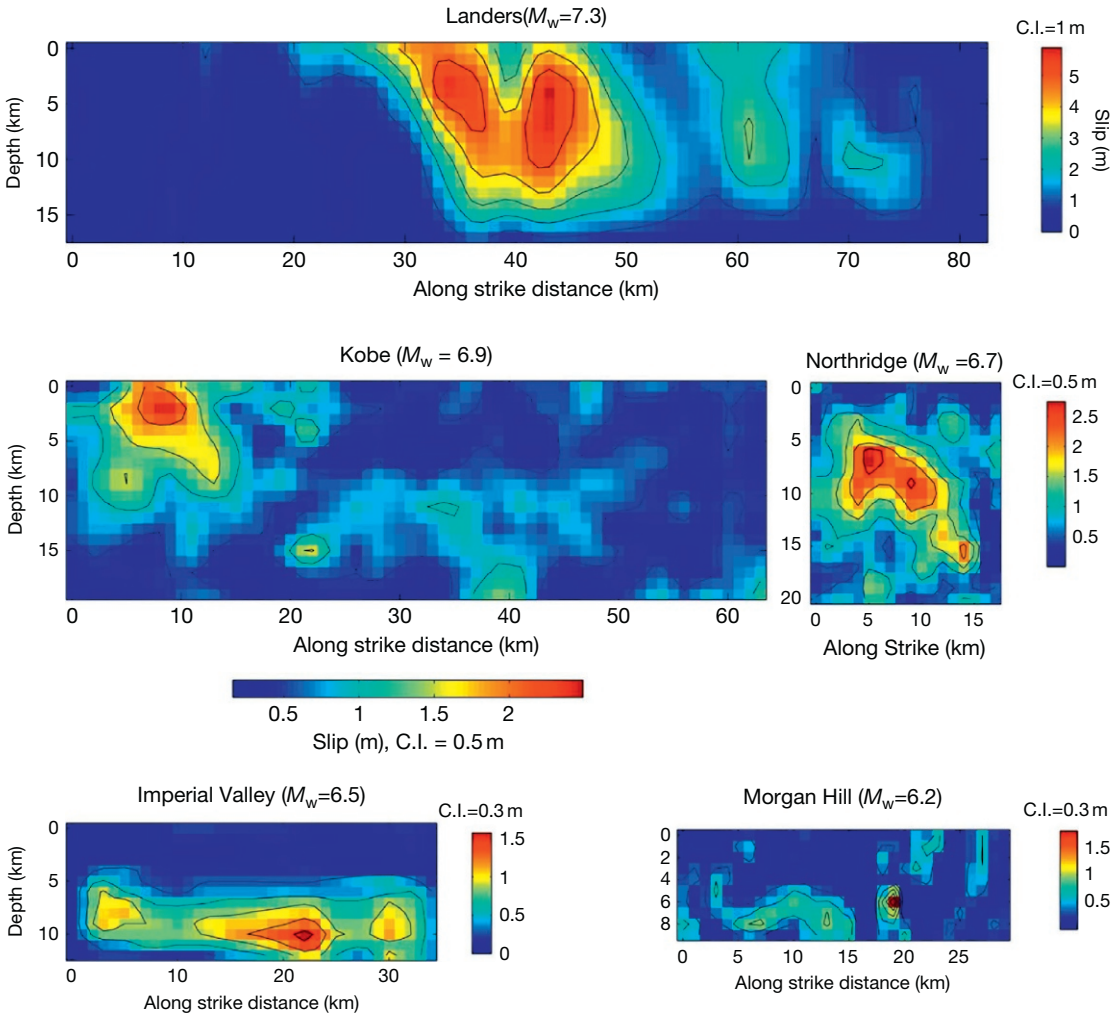


Figure 19 Side view of slip models for five different earthquakes, including the 1992 Landers, California; the 1995 Kobe, Japan; the 1994 Northridge, California; the 1979 Imperial Valley, California; and the 1984 Morgan Hill, California, earthquakes. In each case, slip is strongly variable over the fault surface. Reproduced from Mai PM and Beroza GC (2002) A spatial random-field model to characterize complexity in earthquake slip. *Journal of Geophysical Research* 107(B11): 2308, http://dx.doi.org/10.1029/2001JB000588.

dynamic rupture models does not assume the fault plane a priori but instead allows it to evolve based on a failure criterion for the medium. This is potentially quite important because as rupture speed increases, out-of-plane stresses come to dominate stresses in the direction of rupture and hence favor non-planar faulting.

Dynamic faulting models are typically based on fracture mechanics – an approach that was originally developed for tensile opening-mode cracks. In ideal elastic–brittle crack models, the crack tip separates unslipped from slipping fault surface. Stress in the unslipped region immediately ahead of the crack tip is infinite, but it is an integrable singularity such that the amount of energy expended at the crack tip is finite. Such a singularity is unrealizable in the real Earth, so models have been developed that require a breakdown zone ahead of the crack tip in which shear traction remains finite as the crack tip approaches. For a detailed treatment of the application of fracture mechanics to modeling earthquake dynamics, see Rice (1980) and **Chapter 4.03**.

Dynamic rupture modeling is challenging for several reasons. The need to keep track of stress during rupture means that the entire past history of faulting within the cone of causality is required to calculate the stress on each point on the fault. The need to simultaneously track the large scale of the entire fault rupture and the small-scale heterogeneity near the crack tip presents computational challenges. If grid size is not sufficiently small, then dynamic rupture models can develop an unphysical grid-size dependence. In the presence of material contrast across the fault or of strong velocity weakening, there is the additional complication of numerical instabilities to deal with. Finally, to model multiple earthquake cycles, steady strain accumulation with timescales of centuries and sudden release with timescales of seconds must both be resolved.

A complete discussion of dynamic source modeling is well beyond the scope of this chapter. Recent work in this area has addressed the possibility that material contrast across a fault zone might lead to a direction of preferred rupture (Andrews and Ben-Zion, 1997), whether strong velocity weakening

occurs during seismic faulting (Perrin et al., 1995), the origin of heterogeneity over many earthquake cycles (Lapusta et al., 2000), and attempts to generate dynamic rupture models that emulate the salient properties of kinematic rupture models as in **Figure 20** (Peyrat et al., 2001).

There are several recent trends in dynamic rupture modeling. One is the inclusion of 3-D fault geometry to more accurately portray the likely effects of such heterogeneity on dynamics (Kozdon et al., 2013). Another is the inclusion of off-fault plastic deformation in the presence of large stress concentrations (Dunham et al., 2011), which may be particularly important near the Earth's surface where the confining stress is low (Ma and Andrews, 2010).

4.01.3.6 Volcano Seismology

As with earthquakes, seismology provides a key source of information on how volcanoes work. Unlike earthquakes, eruptions on well-instrumented volcanic systems are routinely predicted and seismological monitoring plays a key role in this capability. Volcanic systems are often highly seismic with literally millions of earthquakes. Volcanoes exhibit a rich range of behaviors that lead to diverse seismic signals.

Figure 21 shows a suite of long-period earthquakes at several volcanoes.

One signal of particular note is volcanic tremor. Volcanic tremor has a range of behaviors, but it often occurs as ground

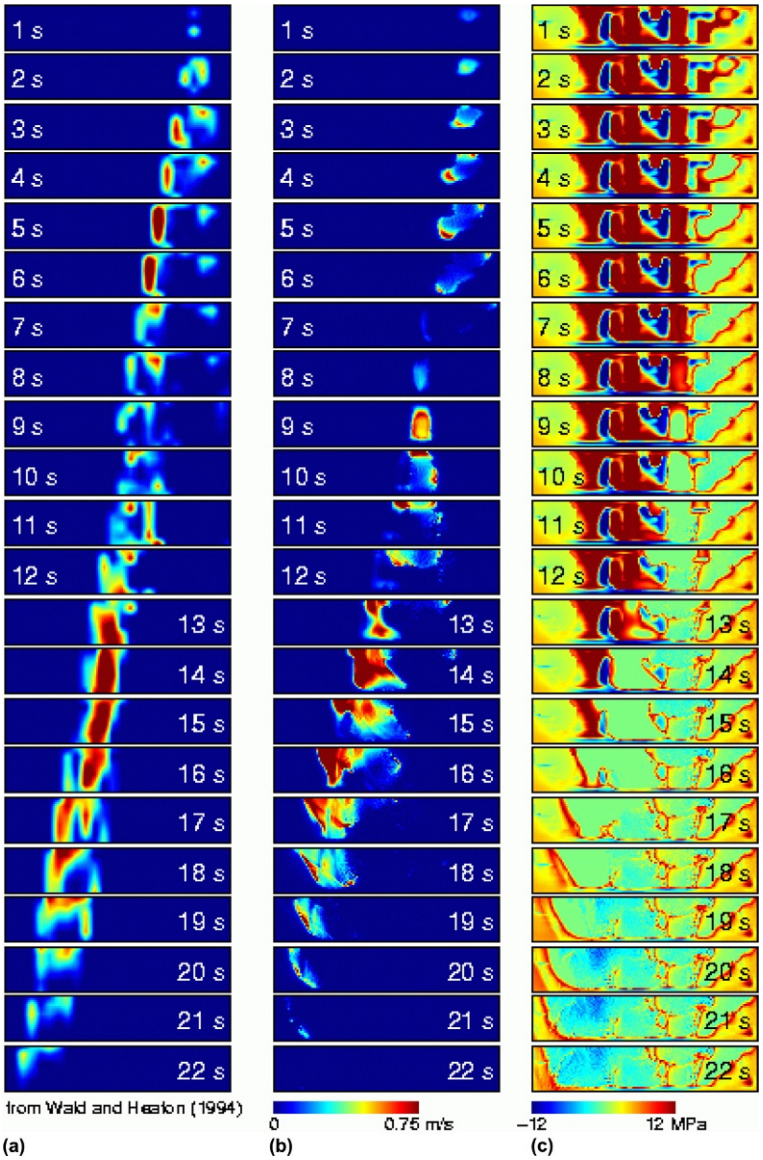


Figure 20 Comparison of kinematic and dynamic rupture models for the 1992 Landers, California, earthquake. Left column shows snapshots of kinematic model from Wald and Heaton (1994), middle column shows snapshots of dynamic rupture models of Peyrat et al. (2001), and right panel shows stress change accompanying dynamic model. Both models fit the long-period strong ground motion data, but predictions of strong ground motion at higher frequencies would vary markedly between the two models. Reproduced from Peyrat S, Olsen KB, and Madariaga R (2001) Dynamic modeling of the 1992 Landers earthquake. *Journal of Geophysical Research* 106: 26467–26482.

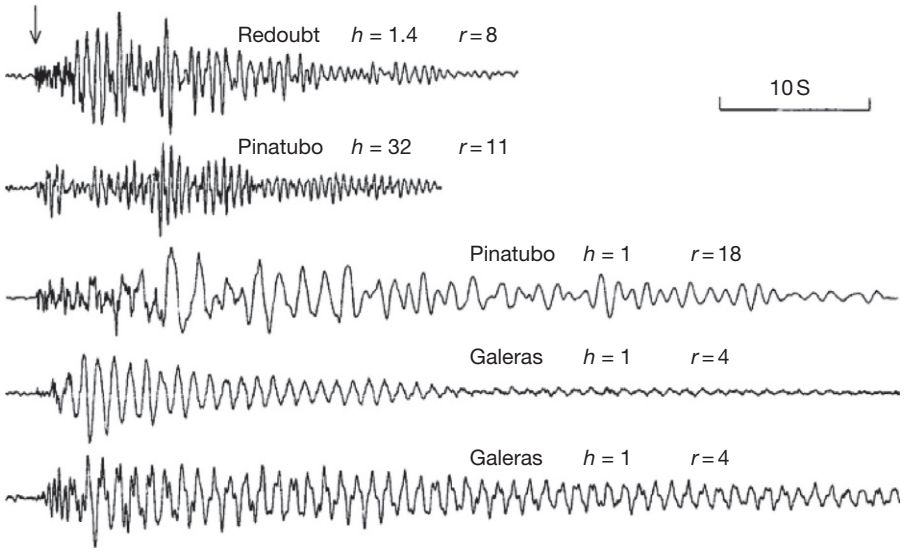


Figure 21 Long-period (LP) earthquakes recorded at Redoubt, Pinatubo, and Galeras volcanoes at source–receiver distance, *r*, and source depth, *h*. These events are thought to reflect pressure fluctuations at depth within a volcano's magmatic system and are useful as a possible precursor to eruption. The lowermost LP earthquake has such a long coda that it verges on harmonic tremor. Reproduced from Chouet B (1996) Long-period volcano seismicity: Its source and use in eruption forecasting. *Nature* 380: 309–316.

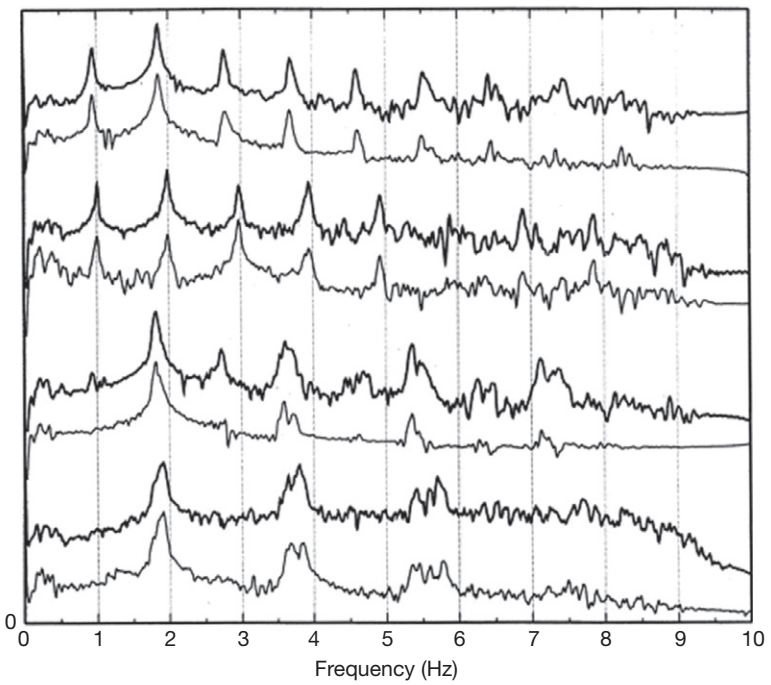


Figure 22 Spectra of velocity seismograms from two seismic stations during four different 30 s time windows during tremor episodes at Arenal volcano, Costa Rica (Hagerty et al., 2000). The peaks at ~1, 3, 5, and 7 Hz are sometimes present and sometimes absent, which is a signature of period doubling.

motion primarily at a discrete set of frequencies. There are several theories for the generation of tremor. A notable example is that due to Julian (1994) who suggested that tremor is generated by a nonlinear interaction coupling between fluid flow and the surrounding elastic material. Among the predictions of this tremor model is that, with increasing flow, tremor that is initially harmonic will undergo a cascade of period doubling as the tremor becomes chaotic. Since that paper was published, period doubling has been observed on several volcanoes including Arenal in Costa Rica (**Figure 22**).

Chapter 4.14 covers volcanic processes and **Chapter 4.15** covers volcano seismology in detail.

4.01.4 Slip Behavior

The nature of slip on earthquake faults varies widely. It can occur suddenly, as in earthquakes. It can also occur quasi-statically, as in fault creep, in which faults slip so slowly that no detectable seismic waves are generated. Documenting the nature of slip behavior, and understanding the reasons for its variation, is a key goal of earthquake seismology.

4.01.4.1 Brittle Failure

The source of nearly all earthquakes is sudden shear slip across a fault. This sudden slip is sometimes referred to as 'brittle' failure. The domain over which brittle failure occurs is dependent on factors such as mineralogy, pressure, and temperature. Because each of these factors varies strongly with depth, synoptic models of earthquake fault behavior usually characterize the nature of slip as a function of depth.

At very shallow depths – such as the upper few kilometers of continental crust in California – geologic materials are weak and unable to store much elastic strain energy. This renders them unable to support sudden slip in earthquakes and is reflected in the depths of the shallowest microearthquakes. On the San Andreas Fault in California, for example, the depths of the shallowest microearthquakes are several kilometers. The San Andreas Fault Observatory at Depth (SAFOD) project targets some of the shallowest seismicity on the creeping section of the fault by piercing the fault at a depth of ~3.2 km. Above this, it is likely that faults slip relatively stably, unless dynamically driven during large earthquake ruptures that originate at greater depths. At somewhat greater depths, rocks have the potential to store elastic strain energy, and the friction on faults is such that they can slip unstably. This depth range is called the seismogenic zone. In California, the seismogenic zone extends to a depth of ~15 km. In oceanic crust, where the mineralogy is characterized by higher melting points, the lower boundary of the seismogenic zone is deeper.

The depth of the lowermost earthquakes, where the mode of deformation changes from brittle to ductile, is sometimes termed the brittle ductile transition or, less commonly, the seismic–aseismic transition. The depth of this transition is controlled by mineralogy, temperature, and pressure. High strain rates, such as those occurring during an earthquake, can result in unstable faulting at greater depths than earthquakes nucleate. High strain rate during aftershock sequences can have a similar effect (Schaff et al., 2002). **Figure 23** shows a schematic interpretation of the depth dependence of geologic processes and their seismological consequences.

Intermediate and deep-focus earthquakes pose an interesting case. The variation with depth of deep earthquakes and the cutoff of the very deepest earthquakes appear to be related to phase changes in the mantle. Because phase changes result in more compact crystal structure, seismologists have long sought a volumetric (isotropic) component to the mechanism of deep-focus earthquakes. They have found, however, that earthquakes at all depths are dominated by deviatoric failure, with deep earthquakes having at most small volumetric components. The fact that deep earthquakes have far fewer aftershocks further distinguishes them from shallow seismicity (Frohlich, 1989). In any case, it is difficult to understand how deep earthquakes can undergo brittle failure at temperatures and pressures where rocks are expected to deform plastically. Possible mechanisms behind deep earthquakes – among them thermal runaway, phase transitions, and transformational faulting – are discussed in **Chapter 4.13**. Regardless of the mechanism, the clear dependence of deep earthquake behavior on the thermal parameter (**Figure 24**) illustrates the important role played by relatively cold subducted lithosphere at great depth.

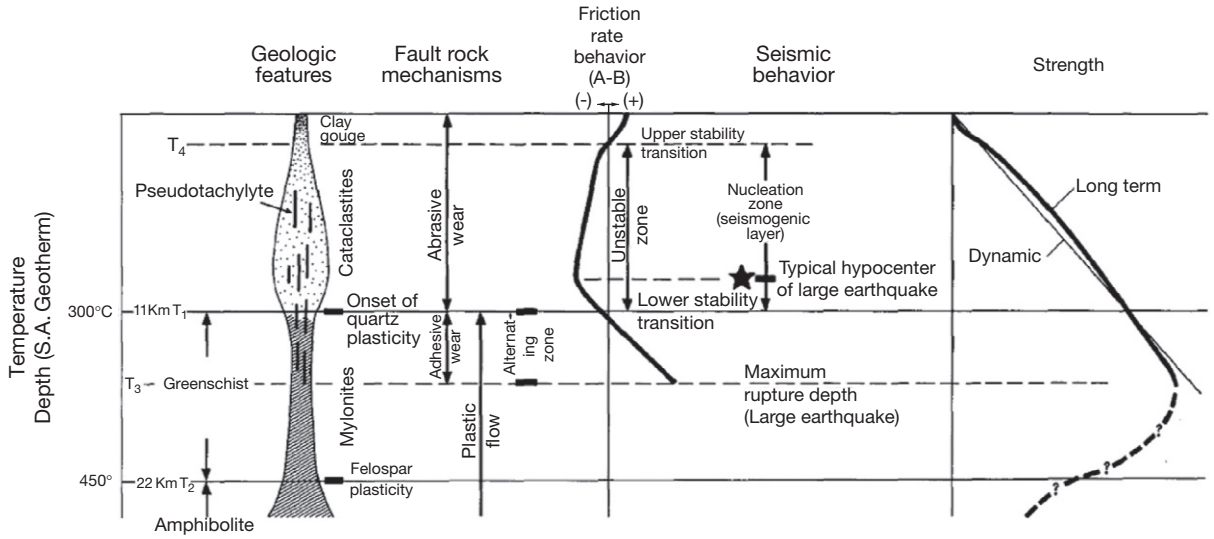


Figure 23 Schematic diagram showing aspects of shear zones and faulting in the Earth. Increasing temperature with depth is shown on the left axis. Geologic manifestation of deformation is shown next to that together with the underlying deformation mechanism. In the middle is frictional behavior, which closely follows the seismogenic zone. Reproduced from Scholz CH (1989) Mechanics of faulting. *Annual Review of Earth and Planetary Sciences* 17: 309–334.

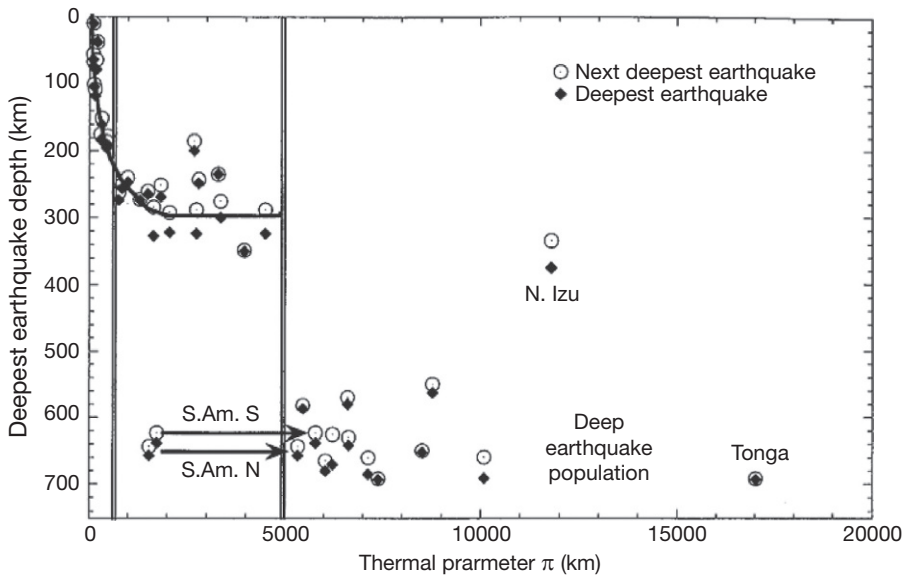


Figure 24 Depth of the deepest earthquake versus the thermal parameter, which is defined by Molnar et al. (1979) as the product of plate age and vertical descent rate, for different subduction zones. A threshold thermal parameter of 5000 appears to separate two domains of behavior. Reproduced from Kirby S, Engdahl ER, and Denlinger R (1996a) Intermediate-depth intraslab earthquakes and arc volcanism as physical expressions of crustal and uppermost mantle metamorphism in subducting slabs. In: Bebout GE, et al. (eds.), *Subduction*: *Top to Bottom*. *Geophysical Monograph* 96: 195–214.2043, http://dx.doi.org/10.1029/2001JB000522, American Geophysical Union; Kirby SH, Stein S, Okal EA, and Rubie DC (1996b) Metastable mantle phase transformations and deep earthquakes in subducting oceanic lithosphere. *Reviews of Geophysics* 34: 261–306.

4.01.4.2 Creep

If slip on a fault occurs slowly enough, then it is possible that no detectable seismic waves will be generated. Such slip is referred to as 'aseismic.' It has long been known that slip occurs aseismically on some faults, such as the San Andreas Fault in central California (Nason and Tocher, 1970). Such slip has come to be called 'creep' and the section of the San Andreas Fault where most of the fault motion occurs aseismically has come to be known as 'the creeping section.' **Figure 25** shows GPS velocity vectors adjacent to the creeping section of the San Andreas Fault in central California, and **Figure 26** shows that creep on this section of the fault at the surface closely matches that inferred over the depth of the seismogenic zone. Once global plate motion models provided predictions about long-term slip rates across plate boundaries, a comparison of earthquake history with seismic slip rates indicated that many plate boundaries have far too little seismic slip (Brune, 1968). Similarly, geodetically measured deformation following significant earthquakes is always observed to be much greater than can be accounted for by slip in aftershocks. The natural inference in both instances is that much of the total fault slip must occur aseismically. An important question about this aseismic slip is whether it is continuous or episodic.

Observations of creep on the San Andreas Fault indicate that creep occurs episodically in a range of sizes that is not unlike the Gutenberg–Richter distribution that applies to ordinary earthquakes (Wesson, 1988). Aseismic transients at depth are sometimes referred to by the oxymoron 'silent earthquakes.' Because they do not radiate seismic waves as ordinary earthquakes do, these events are very difficult to detect; however, they can be detected geodetically using GPS, radar interferometry, creepmeters, tiltmeters, or strainmeters. As these measuring systems become more widespread, the detection of aseismic transients is becoming more common (**Figure 27**). GPS measurements of crustal deformation over subduction zones have led to the discovery of silent earthquakes as large as $M_w \sim 7.5$ (Kostoglodov et al., 2003)

Intermediate between silent earthquakes and ordinary earthquakes are a class of events referred to as 'slow' earthquakes. Slow earthquakes are identified by having low-frequency seismic waves that are much larger than would be predicted based on their strength as measured by seismic waves at high frequencies. They are thought to be slip events in which the rupture velocity and/or slip velocity is significantly lower than ordinary earthquakes. Oceanic transform faults are one environment in which slow earthquakes are common (Okal and Stewart, 1982). They also occur in subduction zones, with some giving rise to devastating tsunamis (Kanamori, 1972). A new class of slow events (**Figure 28**) has recently been discovered in the accretionary prism of the Nankai Trough of Japan (Ito and Obara, 2006). Recently discovered slow events show interesting scaling properties (**Figure 29**).

An earthquake initiates at a point known as the hypocenter. Thereafter, as an earthquake grows, there is a boundary between the part of the fault that is slipping, or has already slipped, and the part of the fault that is as yet unruptured. This boundary is known as the rupture front and the speed at which the rupture front propagates is known as the rupture velocity. The effect of rupture velocity on seismograms was first recognized for the 1952 Kern County, California, earthquake (Gutenberg, 1955) from the asymmetry of the surface wave radiation. Seismological estimates of the average rupture velocity for crustal earthquakes are typically in the range of ~70–90% of the S-wave velocity, or

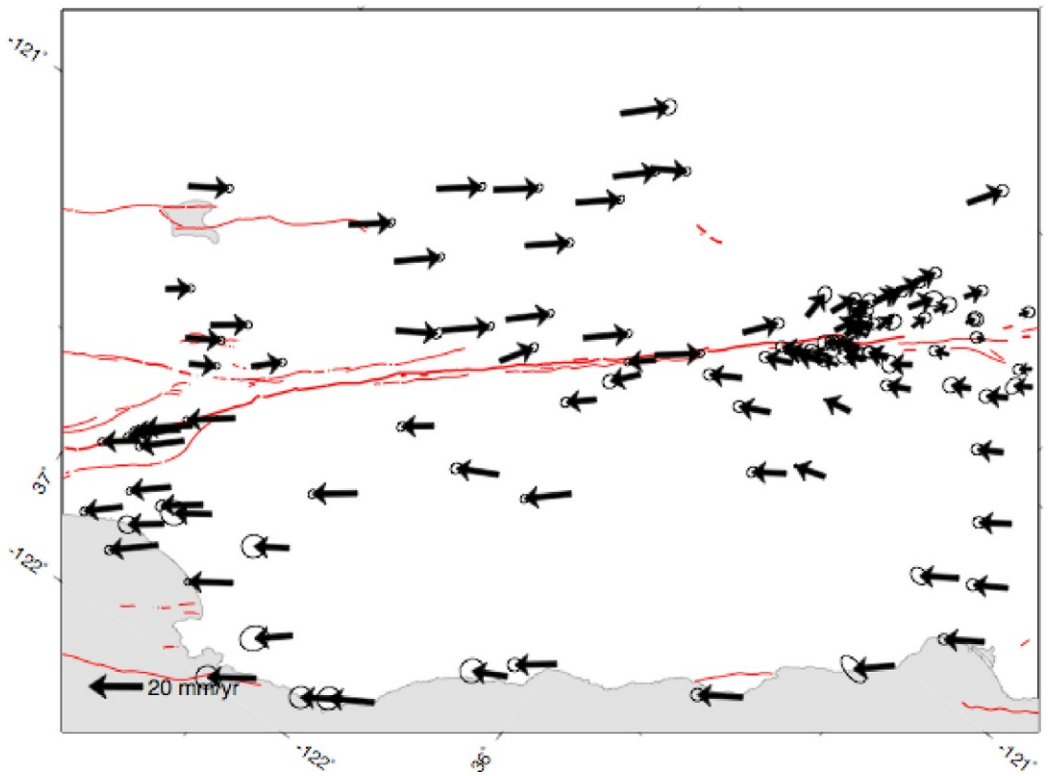


Figure 25 From R. Bürgmann (personal communication) showing faults in red. The San Andreas system of faults crosses the center of the figure. The section between the branching of the fault near San Juan Bautista, on the left of the figure, and Parkfield, the area of dense observations on the right, is known as the creeping section of the San Andreas Fault. GPS velocity vectors shown with arrows and 95% confidence ellipses show nearly block motion of the two sides of the fault (i.e., there is very little velocity gradient with distance away from the fault). Farther to the northwest and to the southeast, the velocity gradient indicates that the fault is locked between large earthquakes.

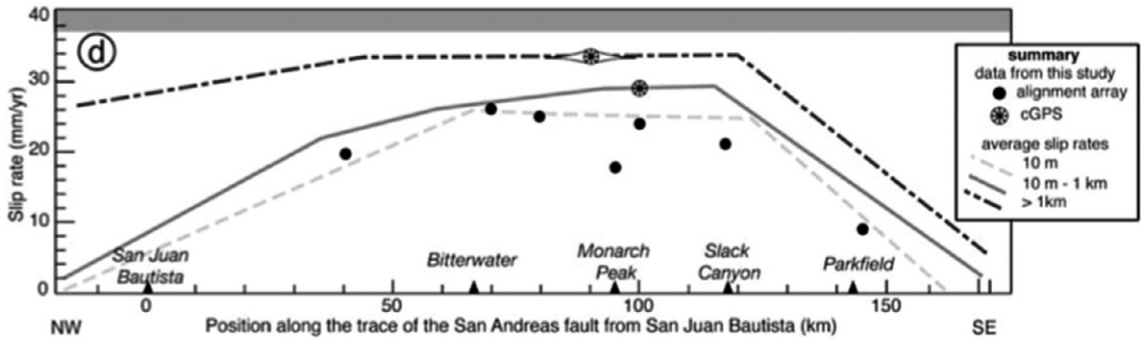


Figure 26 From Titus et al. (2006) showing slip rate parallel to the strike of the San Andreas Fault as a function of distance along the fault from NW (left) to SE (right) as inferred from geodetic measurements. Pale dashed line shows rates determined from measurements within 10 m of the fault and can be considered the surface creep rate. Gray line shows measurements taken within 1 km of the fault and can be considered to represent the slip rate in the upper few kilometers. Dark dashed line shows larger-scale measurements, which include GPS measurements. These estimated slip rates reflect slip to greater depths. The similarity of the curves indicates that creep occurs throughout the depth of the seismogenic zone and the comparison with the model of Argus and Gordon (2001) for the motion of the Pacific plate relative to the Sierra Nevada-Great Valley block indicates that much of that relative motion is taken up by creep at this latitude. The creep rate decays to zero at both the northwest end and the southeast end of the creeping section.

on the order of 2.5–3.0 km s^{-1}. There is no evidence that rupture velocity varies with magnitude, with even microearthquakes having rupture velocities in the same range (Frankel et al., 1986). Rupture velocities that are a large fraction of the shear-wave velocity lead to pronounced directivity in strong ground motion, particularly for shear waves, and lead to strong fault-normal pulses of ground motion in the near field of major faults (Somerville et al., 1997).

From the point of view of earthquake dynamics, it is not clear why earthquakes rupture at these velocities. The simplest models

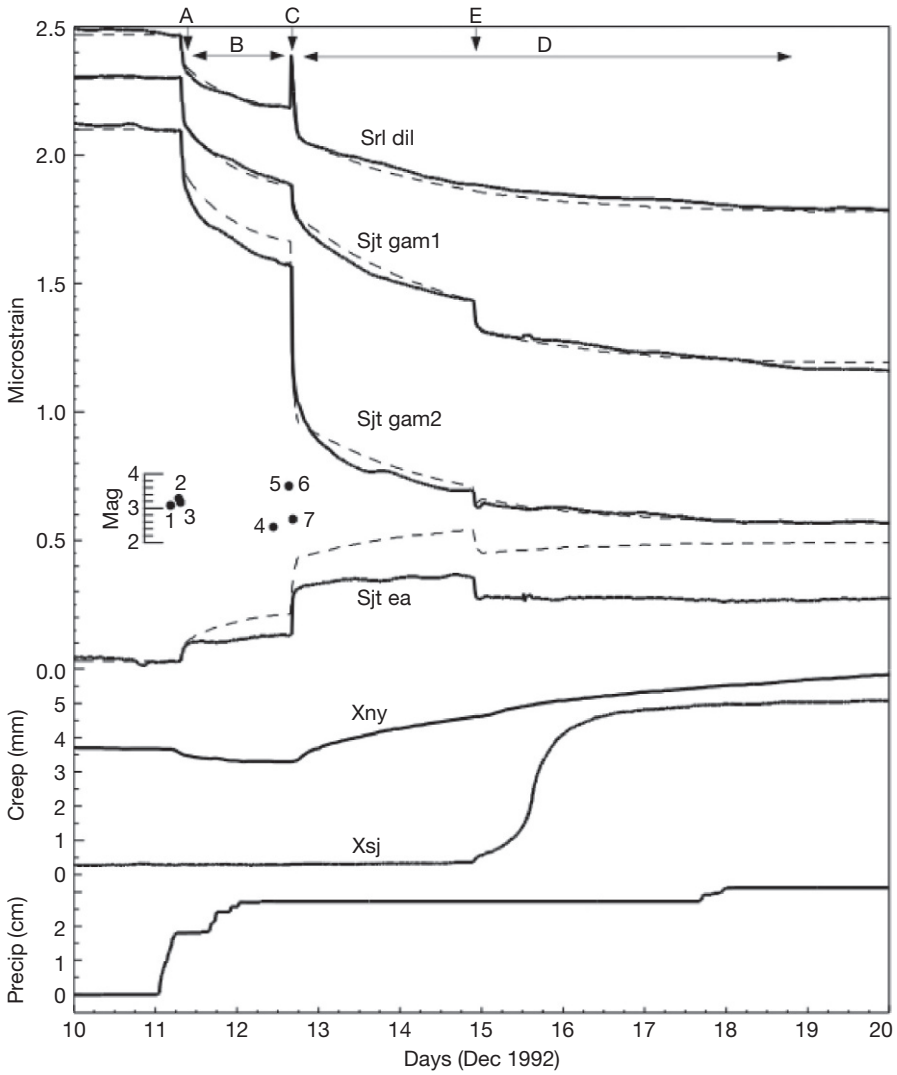


Figure 27 An aseismic transient recorded on the San Andreas Fault near San Juan Bautista. Strain data for stations SRL and SJT, with effects of tides and atmospheric pressure fluctuations removed, are shown as solid lines. Creepmeters (XNY and XSJ) are also shown. Labels A–E denote aseismic transients. Model of strain due to aseismic slip on the San Andreas is shown with dashed lines. Small earthquakes occurred at the same time, but are orders of magnitude too small to explain the observed signals. Reproduced from Linde AT, Gladwin MT, Johnston JS, and Gwyther RL (1996) A slow earthquake sequence on the San Andreas fault. *Nature* 383: 65–69.

based on ideal elastic–brittle fracture theory hold that shear failure ought to accelerate quickly to a limiting velocity that depends on the rupture mode: either the shear-wave velocity for antiplane rupture or the Rayleigh wave velocity (about 92% of the shear-wave velocity) for in-plane rupture (Freund, 1990). Fracture mechanics models also predict that out-of-plane stresses will grow as the limiting velocity is approached, which should promote rupture bifurcation and a lower rupture velocity.

Although sub-Rayleigh rupture velocity is found for most earthquakes, there is increasing evidence that supershear rupture, that is, rupture velocities that exceed the shear-wave velocity, may be an important property of large strike-slip earthquakes. Supershear rupture velocities are expected in slip-weakening rupture models (Andrews, 1976) and have been observed in laboratory fracture experiments (Rosakis et al., 1999). Rupture velocity has apparently exceeded the S-wave velocity for at least several recent large earthquakes with the 1999 Izmit, Turkey (Bouchon et al., 2001); the 2001 Kunlunshan, Tibet (Bouchon and Vallée, 2003); and the 2002 Denali, Alaska (Dunham and Archuleta, 2004) earthquakes all showing characteristics of supershear rupture. If supershear rupture propagation should prove common, it will have important implications for earthquake hazards because the strong ground motion from supershear rupture is qualitatively different from the subshear case (Aagaard and Heaton, 2004). For a full discussion of theoretical, laboratory, and observational aspects of shear rupture velocity, see **Chapter 4.08**.

4.01.4.3 Aseismic Transients and Tremor

Perhaps, the most exciting recent discovery in the field of earthquake seismology is that of large silent earthquakes and

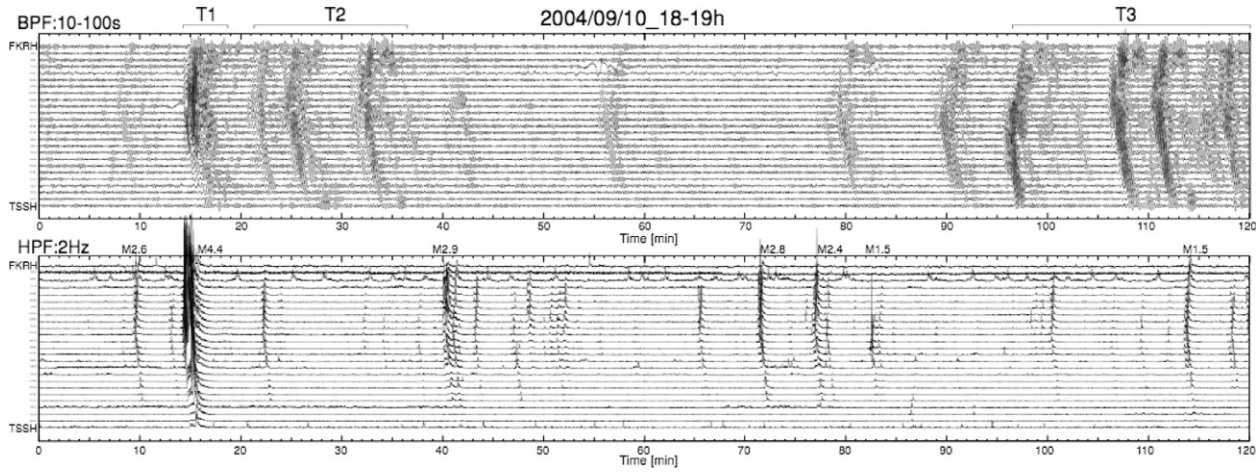


Figure 28 Times series showing very low-frequency earthquakes discovered in the accretionary prism of the Nankai Trough. Two hours of data are shown. Upper panel shows band-passed (10–100 s) ground motion. Lower panel shows rms ground motion for a high-pass filter with corner at 2 Hz. Labels at the top of the lower panel show magnitudes of earthquakes in the JMA catalog. Events in the upper panel at ~25, ~32, ~55, and ~90 min have no corresponding signature at high frequencies and are identified as VLF earthquakes. Reproduced from Ito Y and Obara K (2005) Very low frequency earthquakes excited by the 2004 off Kii peninsula earthquakes: A dynamic deformation process in the large accretionary prism. *Earth Planets and Space* 57: 321–326.

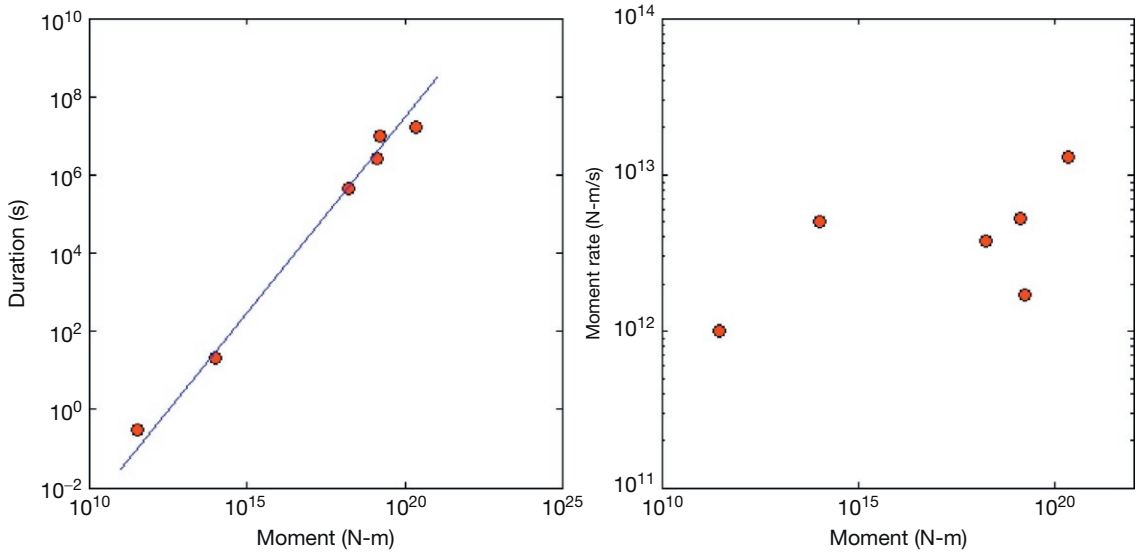


Figure 29 Left panel shows scaling of recently discovered anomalous earthquakes, including (from left to right) low-frequency earthquakes, very low-frequency earthquakes, and slow slip events (see **Chapter 4.17**). The duration scales linearly with seismic moment over 8 orders of magnitude in slow earthquake size. Panel on the right shows the variation of moment rate with changing moment for the same population. The moment rate is nearly constant over 8 orders of magnitude in seismic moment.

accompanying deep, nonvolcanic tremor (**Figure 30**). Deep, nonvolcanic tremor was discovered in Japan in the parts of the subduction zone where the Philippine Sea plate is subducting (Obara, 2002). It occurs at depths where dehydration reactions are thought to liberate water from the subducting plate (Seno and Yamasaki, 2003), and hence, it is natural to attribute the genesis of tremor to the movement of fluids at depth. Fluid generation of tremor is also thought to be operative in volcanic tremor that is observed in active magmatic systems (Chouet, 1988; Julian, 1994). Large silent earthquakes were discovered in Cascadia by Dragert et al. (2001) and subsequently found by Miller et al. (2002) to be occurring periodically with a recurrence interval of 14 ± 1 months. The remarkable regularity of recurrence of these silent earthquakes stands in apparent stark contrast to the more variable recurrence of more typical large earthquakes.

A near-perfect correlation between tremor and silent earthquakes in Cascadia was discovered by Rogers and Dragert (2003) and is illustrated in **Figure 31**. A correlation between the two phenomena was subsequently documented in Japan as well (Obara et al., 2004), where silent earthquake recurrence is

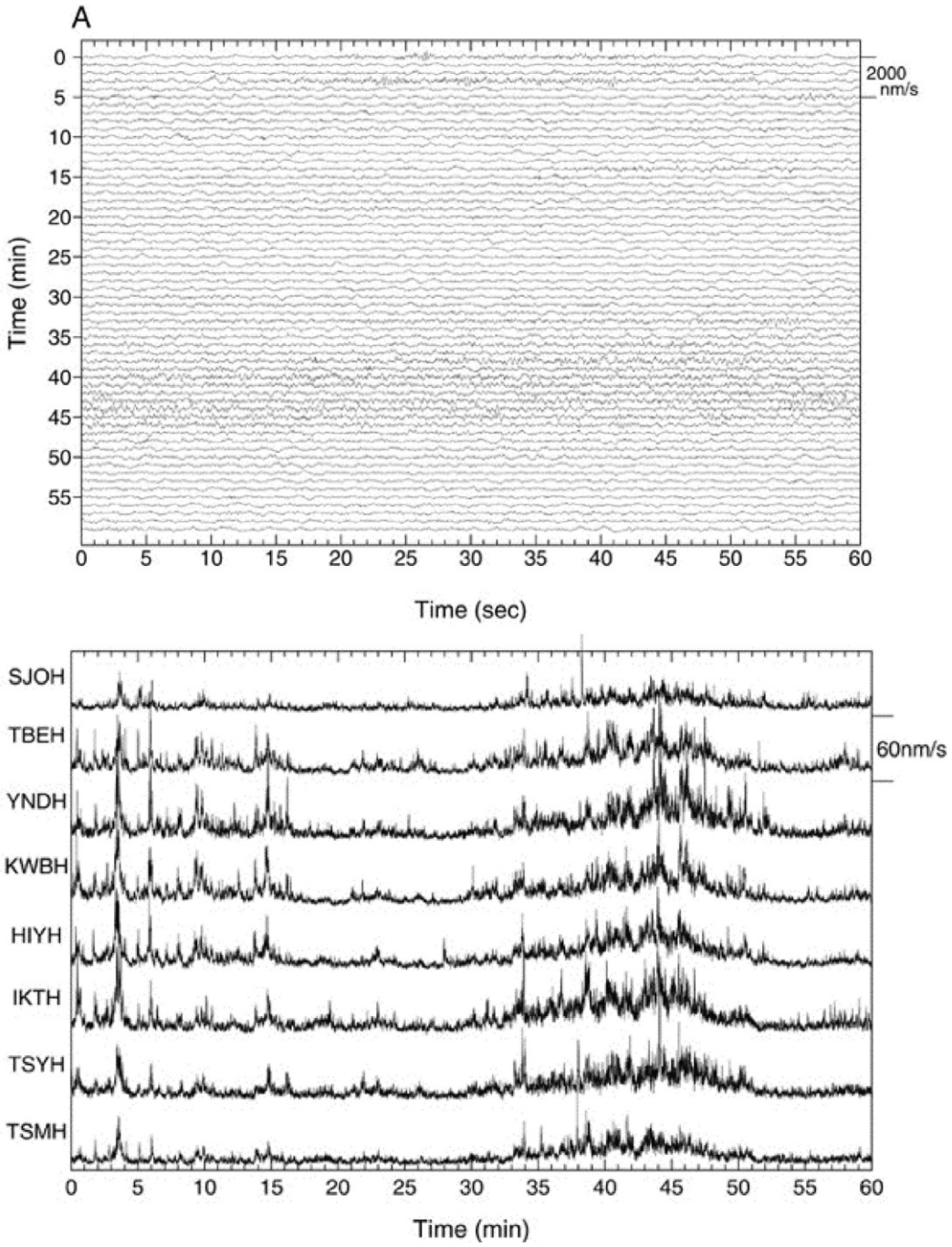


Figure 30 An example of signals from deep nonvolcanic tremor as observed by Obara (2002). The upper panel shows ground motion at a single station of the Hi-net borehole array of high-sensitivity/low-noise seismic network. The tremor is strongest from 40 to 47 min and is characterized by a chaotic signature with no clear, impulsive arrivals. Such signals could be attributable to a range of sources, but their similarity at widely spaced stations and the high phase velocity between them indicates that the signals have a common source, originating deep within the Earth.

less regular, but on the order of several months in eastern Shikoku. The two phenomena are so closely intertwined that they are commonly referred to with an acronym ETS, for episodic tremor and slip. Longer-term aseismic transients, lasting years, have been documented in several areas of Japan including the Bungo Channel (Hirose and Obara, 2005) and the Tokai region (Ozawa et al., 2005). In each of these cases, the aseismic transients appear to arise due to aseismic slip on the plate interface just below the locked seismogenic zone. Given this location, they are certain to accelerate the loading of the source region of large megathrust earthquakes and hence should be capable of triggering such events.

The location of the tremor source at depths at which metamorphic dehydration reactions are expected and the fact that

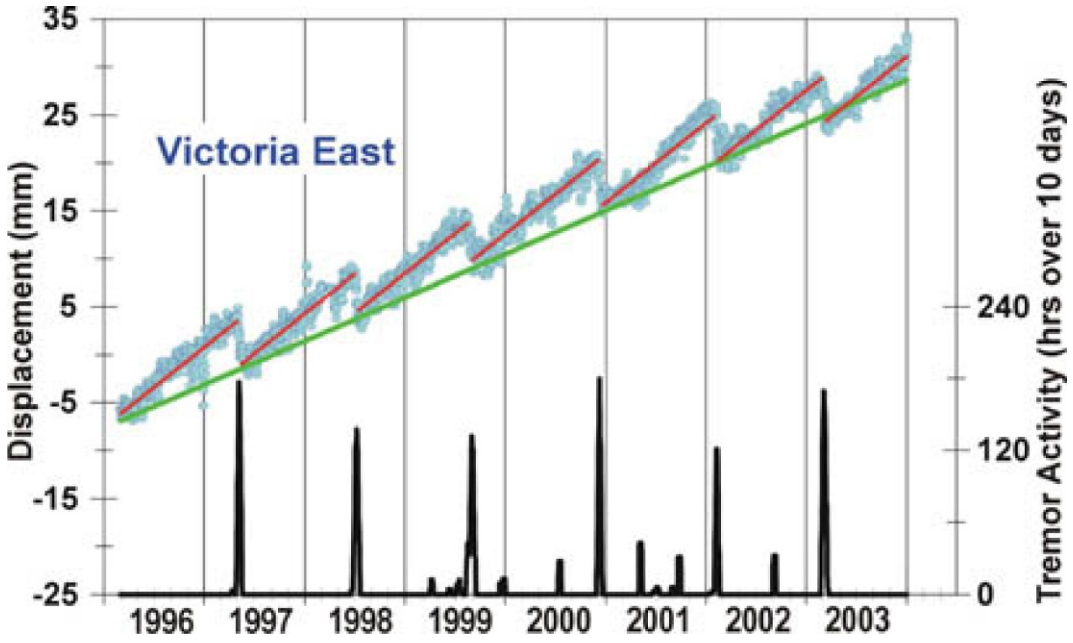


Figure 31 Comparison of slip and tremor activity for Victoria, British Columbia, Canada (http://seismescanada.rncan.gc.ca/pprs/research/ETS_e.php). Small blue circles show position of a GPS site near Victoria with respect to the interior of North America. Green line shows long-term eastward motion of the site. Red sawtooth line shows motion and includes reversals in motion approximately every 14 months. These and displacements observed at other GPS stations have been modeled as aseismic slip on the plate interface at depth. The bottom black graph shows the number of hours of tremor activity observed on southern Vancouver Island. The correlation between episodic tremor and aseismic slip is remarkable and has led to the two phenomena being referred to as episodic tremor and slip (ETS).

tremor is observed for Philippine Sea plate subduction, but not Pacific plate subduction, in Japan both point to a key role for fluids in the process (Julian, 2002). Fluids either may couple directly into ground motion as they are thought to do in the case of volcanic tremor or may act to trigger both phenomena. The discovery by Shelly et al. (2006) that low-frequency earthquakes that occur during tremor episodes represent slip on the plate interface indicates that at least part of the tremor signal arises from shear slip. Subsequent work (Ide et al., 2007a; Royer and Bostock, 2013; Shelly et al., 2007; Wech and Creager, 2007) supports the notion that the entire tremor signal is generated as shear slip on faults and that tremor is part of a family of slow earthquakes (Ide et al., 2007b). For more details on aseismic transients, see **Chapter 4.17**.

4.01.5 Physics of the Earthquake Source

In this introductory chapter, we have focused on the phenomenology of the earthquake source as it is studied using seismology. For the most part, we have not focused on the physical mechanisms underlying earthquake behavior. There is much about these physical mechanisms that we have yet to understand. In this section, we attempt to tie together the physics behind the observations.

4.01.5.1 Friction

Most earthquakes occur on preexisting faults. The friction acting across these faults provides the resistance to keep them from slipping in between earthquakes. Laboratory measurements of friction indicate that for most rocks at seismogenic depth, the coefficient of friction should be ~0.6 (Byerleee, 1978). For earthquakes to propagate unstably, however, friction has to drop from a static level to some lower sliding level, as in the slip-weakening model (see next section). Even slip-weakening behavior, however, does not allow for faults to recover their frictional strength once slip has occurred, so realistic faults require a more complex frictional behavior.

More complex frictional behavior is, in fact, observed in laboratory experiments, and these second-order frictional effects may control much of fault behavior in the Earth. Static friction is not constant, but depends on the history of slip on the interface, with friction increasing as the logarithm of time of stationary contact in laboratory experiments (Dieterich, 1972). A related property is that dynamic sliding friction depends logarithmically on the slip rate. The sign of this dependence determines whether the dependence is weakening or strengthening with increasing slip rate (Scholz et al., 1972). A third property is that when the slip rate on a sliding surface changes suddenly from one steady rate to another, the coefficient of friction evolves to a new steady-state value after a critical amount of slip, D_c, has occurred. D_c is thought to represent the slip required to renew the true contact area between two rough surfaces (Dieterich, 1978). These three properties – the time dependence of static friction, the rate dependence of dynamic friction, and the evolution of sliding friction with slip – have been unified in a theory of rate- and state-dependent friction laws that describe the evolution of friction (Dieterich, 1979a,b; Ruina, 1983).

An important implication of these laws for earthquakes is that the effect of velocity arises from the combination of two

effects – the direct velocity effect and the evolution of friction with slip – that have opposite signs. Whether slip can occur unstably or must occur stably depends critically on their relative size. Rocks with steady-state velocity strengthening will always slip stably, while rocks with steady-state velocity weakening can slip unstably. The transition from steady-state velocity-weakening to steady-state velocity-strengthening behavior is thought to control the lower (Tse and Rice, 1986) and upper (Marone and Scholz, 1988) depth extent of the seismogenic zone. More on this topic can be found on **Chapters 4.06** and **4.04**.

4.01.5.2 Energy Budget

In the time between large earthquakes, seismogenic faults remain frictionally locked. As plate motion continues, the crust around a fault will gradually distort and store energy in the form of gravitational energy and elastic strain energy. When the fault starts to slip, this gradually accumulated potential energy is suddenly converted to different forms. Some of it is used to overcome the initial resistance to shear failure. This component of the energy is referred to as the fracture energy. As slip continues during the earthquake at some lower frictional level, energy is converted to heat due to the work done against the resisting frictional force. Another term in the earthquake energy budget is the component that is radiated in the form of seismic waves. This is the seismic energy referred to in **Sections 4.01.2.1** and **4.01.3.4**. Seismic energy is gradually dissipated and converted to heat elsewhere due to anelastic wave propagation effects. The energy balance can be expressed using a simple equality (Kostrov and Das, 1988) involving the radiated seismic energy, E_s; the change in potential energy, ΔE_p; the energy expended extending the crack, ΔE_γ; and the energy expended overcoming friction, ΔE_f:

$$E_s + \Delta E_p + \Delta E_\gamma + \Delta E_f = 0$$

Of the terms in this equation, we have discussed the radiated seismic energy, a quantity that can be measured using seismic waves. The seismic energy can (with considerable uncertainty) be determined from seismological observations, and it is natural to ask what constraints the seismic energy places on the other terms in the energy balance.

A further observational constraint comes from measurements of the slip. Together with the friction, the slip determines how much energy is dissipated as heat on the fault. This is not as simple as it sounds, however, as it is unrealistic to assume that sliding friction is constant. Rate- and state-variable friction laws indicate otherwise, and there are other mechanisms, discussed in the subsequent section, that are expected to introduce much larger dynamic variations in fault strength.

The fracture energy has been estimated for a number of earthquakes. The total fracture energy is proportional to the faulted area, so if these measurements represent fracture energy on a planar fault, they would only be of consequence for very small earthquakes. Seismological observations of the shear fracture energy are quite large, on the order of 10^6 J m^{-2}. This is orders of magnitude larger than what is measured in the laboratory (Wong, 1982). One interpretation of this discrepancy is that the dissipation of energy being measured includes energy lost due to various forms of yielding near the rupture front (Andrews, 2005). It has been suggested that finely pulverized rock observed in the near field of the San Andreas Fault and the tremendous surface area created in that process might be large enough to account for seismologically inferred fracture energies.

In the slip-weakening model (Ida, 1972), the resistance of the fault to motion is slip-controlled and decreases from some maximum when slip first begins, to a lower sliding frictional level once the slip-weakening distance is attained. Thereafter, slip can continue at the sliding frictional level of resistance. Seismologists have attempted to place constraints on the slip-weakening displacement. They have found that it is quite large in major earthquakes, reaching a value of $\sim$0.5–1 m for the Kobe earthquake (Ide and Takeo, 1997). Here again, the inferred values are orders of magnitude larger than would be inferred by extrapolating measurements from the laboratory. At least some of this discrepancy may result from a lack of adequate resolution of the slip velocity function (Guatteri and Spudich, 2000).

Other informations on the energy balance may come from magnitude dependence of the radiated seismic energy. If the true fracture energy is only an important component of the energy budget for very small earthquakes, then one would expect to see more seismic energy radiated per unit fault slip for larger earthquakes. Seismic moment is more easily calculated than fault slip, so seismologists typically consider changes in the ratio of seismic energy to seismic moment, that is, E_s/M_0. A possible source of a change in E_s/M_0 comes from the possibility that the resistance to fault motion may drop dramatically with increasing fault slip (Kanamori and Heaton, 2000) as discussed in the next section. Another interpretation is offered by Abercrombie and Rice (2005), who replaced the slip-weakening distance with a slip-weakening function for which the resistance to fault slip continues to decrease with increasing fault slip. As with more dramatic weakening mechanisms, E_s/M_0 is predicted to increase as an earthquake grows larger.

So what do the data indicate? There is a lot of scatter in the measurements, but representative values for large earthquakes are on the order of 5×10^{-5}. Some observations of the magnitude dependence of seismic energy suggest that the ratio of seismic energy to seismic moment increases as a function of earthquake size. That is, earthquakes become more efficient at radiating seismic waves as they grow larger. Measurements from a number of studies of earthquakes over a wide range of earthquake size indicate that E_s/M_0 is approximately a factor of 10 larger for M_w 7 earthquakes than it is for M_w 2 earthquakes (Kanamori and Heaton, 2000). Ide and Beroza (2001) found that many published E_s/M_0 measurements could be biased to low values for small earthquakes because of biases introduced by inadequate corrections for path effects or the limited instrumental bandwidth (**Figure 32**). Thus, whether or not the observed scaling of E_s/M_0 with earthquake size is a true property of earthquakes remains an open question.

4.01.5.3 Microscopic Processes

There are many processes that occur at the microscopic scale that may have no direct seismological signature, yet still have a profound effect on and may ultimately control macroscopic

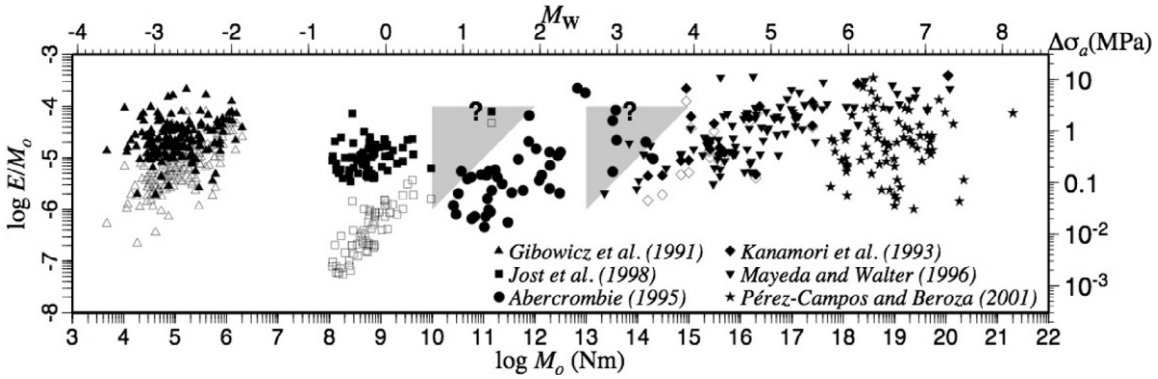


Figure 32 Plot of E_s/M_0 versus M_0 from several studies. Open symbols show original measurements, while filled symbols show values once seismic energy possibly missing due to limited bandwidth is restored. Gray triangles show areas where earthquakes are excluded from two studies due to limited bandwidth. Reproduced from Ide S and Beroza GC (2001) Does apparent stress vary with earthquake size? *Geophysical Research Letters* 28: 3349–3352

earthquake behavior. In this section, we briefly review several such phenomena.

4.01.5.3.1 Hydrologic processes

An important unknown in earthquake science is the importance of pore fluids on the faulting process. Fluids support normal stress, but do not support shear stress. Thus, one way in which pore fluids can effect faulting is by changing the effective normal stress. Briefly, the criterion for slip under the Coulomb failure criterion (e.g., Harris, 1998) states that fault slip will occur when the shear stress, τ, acting across the fault exceeds the sum of the cohesion, C, and the effective normal stress multiplied by the coefficient of friction, μ:

$$\tau > \mu(\sigma - P) + C$$

where the term in the parentheses is referred to as the effective normal stress and is the difference between the fault-normal stress, σ, and the pore pressure, P. The cohesion is usually assumed to be negligible. By this criterion, one way in which earthquakes can be triggered is by an increase in the pore pressure, P.

This effect has been cited by Hubbert and Rubey (1959) as a possible explanation for the movement of large, nearly horizontal thrust faults due to the large frictional resistance expected from the considerable overburden. Fluids have been shown capable of artificially triggering earthquakes in controlled experiments (Healy et al., 1968; Raleigh et al., 1976); thus, it is reasonable to expect that they might play a similar role for tectonic earthquakes. Both the lack of a heat flow anomaly (Brune et al., 1969) and the nearly fault-normal stress orientation of some major fault zones (Zoback et al., 1987) could be understood if the effective normal stress is low, due to high pore fluid pressure, such that the fault slips at low shear stress. There are independent evidences that fluids and faults interact (Sibson, 1981) and that high pore pressures occur in fault zones (**Figure 33**), at least at shallow depths (Unsworth et al., 1997). Dehydration embrittlement, in which phase changes that expel water decrease the effective normal stress in subducting lithosphere, is cited as a possible enabling mechanism for intermediate-depth earthquakes (Kirby et al., 1996a,b).

Another way in which fluids can influence faulting is through poroelastic effects. In a poroelastic medium, pore pressure will decrease where the mean normal stress decreases and increase where the mean normal stress increases. The immediate change in pore pressure, ΔP, for what is known as undrained conditions can be related to the change in the mean normal stress, $\Delta\sigma_{kk}$, by the equation

$$\Delta P = B\Delta\sigma_{kk}$$

where Skempton's coefficient, B, a function of elastic moduli for drained and undrained conditions, is equal to one for fluid-saturated soils and ranges between 0.5 and 0.9 for a range of rock types (Rice and Cleary, 1976). With time, this undrained state evolves to the drained state as fluids flow and reequilibrate in response to pore pressure gradients.

Pore fluids can trigger earthquakes by inducing shear stress through a process referred to by Booker (1974) as consolidation. Consolidation in this usage refers to the process in which, in places where the mean normal stress and pore pressure increase, normal stress and pore pressure then gradually decrease as pore fluids migrate away from areas of high pore pressure. Where the mean normal stress and the pore pressure decrease, the mean normal stress will decrease further as pore fluids migrate into areas of lower pore pressure. The net effect is to couple normal and shear stress through pore fluid migration. *Booker* proposed this as a candidate mechanism for the gradual decay described in Omori's law of aftershock decay.

Yet another way in which fluids can influence earthquakes and faulting is through chemical effects and stress corrosion (Scholz, 1968). It is also worth noting that other fluids, notably magma and carbon dioxide (Miller et al., 2004), can have similar effects. For more details on the involvement of fluids in faulting, see **Chapter 4.12**.

4.01.5.3.2 Melting and flash heating

Rapid shear slip during earthquake faulting will lead to shear heating. If this heating is large enough, it can lead to very large changes in temperature and, possibly, to melting on the fault (Jeffreys, 1942; McKenzie and Brune, 1972). There are several variables that control whether or not melting occurs and, if it does, how extensive it is. The heat generated per unit area of the

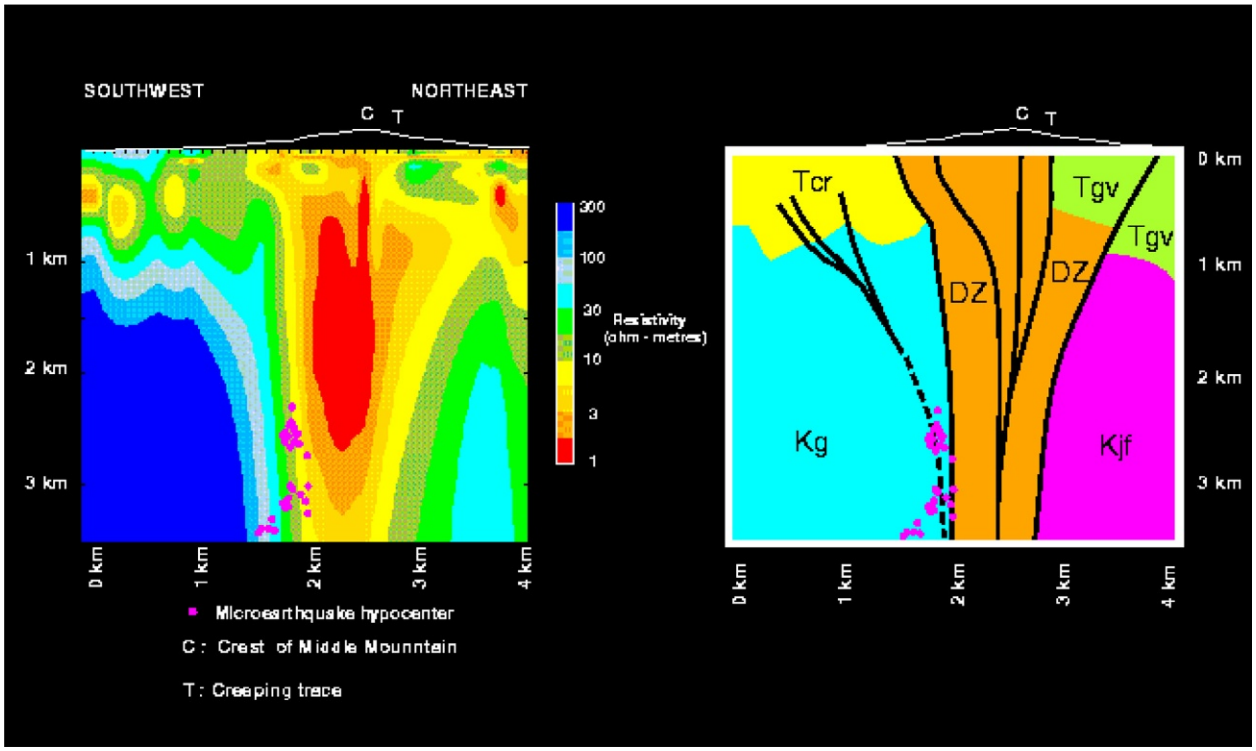


Figure 33 A cross section of imaged resistivity across the San Andreas Fault at Parkfield California (Unsworth et al., 19l7). Resistivity (in Ωm) is shown on the left panel, with a corresponding geologic interpretation on the right panel. The areas of low resistivity (orange red) are interpreted as a fluid-saturated damaged zone (DZ) centered on the fault.

fault will be the product of the slip and the frictional resistance, so the slip and sliding friction will in part determine the conditions for possible melt generation. Melt may begin to form microscopically in the immediate vicinity of frictional contact, leading to initial loss of strength in a process known as 'flash heating' (Rice, 2006).

In addition, any real fault zone will be a shear zone of finite thickness. The thicker the shear zone, the more distributed heat production will be, and all other things being equal, the more slip required to initiate melting. The thermal conductivity will control how quickly heat diffuses away from the fault. Thermal conductivity of rocks is low enough that this will not likely be an important effect, unless either the slip rate is very low or pore fluids are present. The effect of melt on friction is not straightforward. Laboratory experiments suggest that the initial effect might be to increase fault strength due to viscous effects, but this effect might be diminished at high stress. In any case, further slip will lead to more heating and a reduction in viscosity and hence strength (Fialko and Khazan, 2005). A detailed treatment of temperature effects during faulting can be found in **Chapter 4.03**.

4.01.5.3.3 Thermal pressurization

The same frictional heating discussed in the previous section, in the presence of pore fluids, can lead to thermal pressurization of those fluids well before the onset of melting. Thermal pressurization can have a strong effect on faulting by reducing the effective normal stress. The absence of geologic evidence for melting near fault zones led Sibson (1973) to propose thermal pressurization as a mechanism for reducing dynamic fault strength. For undrained conditions, a temperature change of ΔT leads to a pressure change, ΔP of, $\Delta P = \Lambda \Delta T$, where Λ is the thermal expansion coefficient of water. Representative values of Λ for upper crustal conditions are 0.78–0.93 MPa/°C (Rempel and Rice, 2006). Thus, thermal pressurization is potentially a very large and efficient agent of fault weakening. Thermal pressurization was explored as a potential explanation for the lack of a heat flow anomaly on the San Andreas Fault by Lachenbruch and Sass (1980). For it to operate, fluids must be present and fault-zone permeability has to be low enough that the excess pressure does not diffuse into the host rock. Measurements of the permeability of fault zones and their surroundings are very low (Wibberley and Shimamoto, 2003); however, any dilatancy effects during rapid slip could rapidly increase that permeability and act to short circuit this weakening mechanism. More on this mechanism can be found in **Chapter 4.03**.

4.01.5.3.4 Silica gel formation

Another candidate weakening mechanism is the formation of silica gels (Goldsby and Tullis, 2002). In this scenario, sliding in the presence of water leads to the formation of a gel that facilitates sliding at low friction. Evidence for this comes primarily from the laboratory where sliding at fault slip speeds appropriate to earthquakes can lead to dramatic reductions in friction and for which flow textures are preserved as amorphous solids after sliding ceases (Di Toro et al., 2004). There is evidence for silica gels having formed on faults within the

Earth as well (Kirkpatrick et al., 2013). The laboratory data suggest that this mechanism, if important for large earthquakes, would only seem likely to operate in rocks of high silica content.

4.01.5.4 Fault-Zone Structure

Detailed geologic observations place important constraints on the physical processes that operate during faulting in the Earth. The geometric complexity of faults has a first-order effect on the faulting process. Geologic observations also provide constraints on the nature of the deforming or damaged zone in the direction normal to the fault. These observations are crucial for understanding the mechanics of faulting and bear directly on some of the most important problems in earthquake physics.

Seismological constraints on fault-zone structure are of much lower resolution, but they have the advantage of measuring properties of faults that are actively deforming and not exhumed. Earlier in this chapter, we outlined some of the recent progress on the fine structure of faults as revealed by precise earthquake locations (**Figures 11** and **12**). Precise earthquake locations bear directly on questions such as how wide the fault zone is at seismogenic depths and whether or not fault trace complexities observed at the surface reflect complexities in fault geometry at depth.

Another important constraint on the structure of fault zones comes from fault-zone guided waves (FZGWs). Because faults move geologic units past one another, they tend to be planes that separate materials with different properties. Moreover, the mechanical, chemical, and/or hydrologic alteration of materials within deforming fault zones leads to them being a locus of low seismic velocities. Variations in material properties either within or across fault zones lead to strong effects on seismic wave propagation. Waves that are trapped within the low-velocity material within several hundred meters of active fault zones are known as FZGWs. These are analogous to Love waves in vertically layered media in that they usually consist of critically reflected S-waves within the low-velocity material. FZGWs have been observed in the subduction zone in Japan (Fukao et al., 1983), continental normal faults (Leary et al., 1987), and continental strike-slip environments (e.g., Li et al., 1990). Waves that refract horizontally due to a large contrast in velocity across a fault zone are known as fault-zone head waves. These have also been observed in diverse environments, including both subduction zones (Fukao et al., 1983) and continental strike-slip faults (e.g., McNally and McEvilly, 1977).

Fault-zone waves are very sensitive to the details of fault-zone structure and have the potential to place important constraints on the structure of active fault zones. If the geometric regularity of the waveguide is disrupted, then such waves may not propagate or may be strongly altered. For this reason, they can be used to explore fault segmentation. An outstanding question regarding FZGW is just how deep the waveguide extends. Some investigators find that it spans the entire depth of the seismogenic zone (Li et al., 2000), while others argue that it may only extend through the upper few kilometers (Ben-Zion et al., 2003). The recent observation of FZGW in the SAFOD borehole through the San Andreas Fault suggests that the fault-zone waveguide extends to seismogenic depths (**Figure 34**).

Other important questions regarding the fault-zone waveguide are to what extent it is a permanent feature and how much it is a transient that follows a large earthquake. Following the Landers earthquake, there is a clear and strong postseismic healing, or recovery of low velocities, indicating that at least some of the low-velocity zone is transient (Li et al., 1998).

The relationship between low-velocity fault zones and the mechanical behavior of faults in earthquakes is important for earthquake faulting. Observations of surface deformation obtained by satellite radar interferometry (Fialko et al., 2002) following the 2001 Hector Mine, California, earthquake indicate that nearby fault zones are characterized by relatively compliant zones that are on the order a kilometer wide. The effect of relatively low-strength materials is an important question that is just now being addressed.

4.01.5.5 Borehole Observations, Fault-Zone Drilling, and Seismicity in Mines

Another approach to the study of earthquakes and faulting is to make observations below the Earth's surface. This can take the

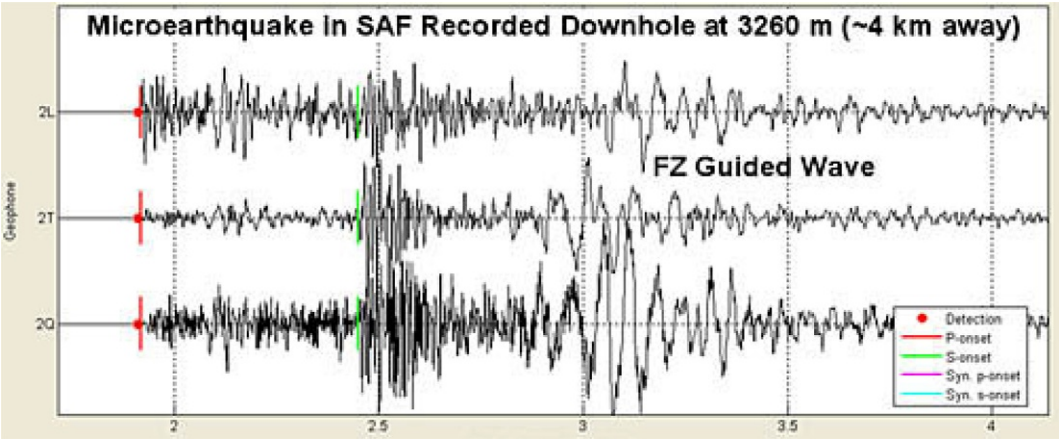


Figure 34 Fault-zone guided waves observed at 2.7 km depth in the SAFOD borehole from a magnitude 1.3 earthquake. The presence of these waves at this depth is diagnostic of a 200 m wide low-velocity zone associated with the San Andreas Fault. Reproduced from http://www.earthscope.org/news/announcements/2006/FaultZone2006/seismogram-big.jpg.

form of recording earthquakes with seismometers deployed in boreholes, direct drilling into active faults, or studying seismicity in deep mines. This approach has a number of distinct advantages.

Borehole seismic observations allow earthquake scientists to make observations below the highly attenuating and strongly scattering layers of the Earth's surface. Borehole measurements have had an important impact on seismology as they allow a relatively unobscured view of seismic waves as they are radiated from the earthquake source (see the section on stress within the Earth in the succeeding text). The Hi-net seismic network in Japan capitalizes on this idea as it uses sensitive borehole instruments to monitor the entire Japanese archipelago. It led directly to the discovery of deep, nonvolcanic tremor (Obara, 2002). The idea of making measurements in the very near field of a fault at depth is a key element of the ongoing SAFOD project.

Deep drilling allows direct sampling of active faults *in situ*, which is a tremendous advantage. It also allows scientists to make stress measurements at depth near active faults. However, fault-zone drilling is a logistically challenging and expensive undertaking. These disadvantages of fault-zone drilling have prevented its wider application. The expense of drilling rises rapidly with increasing depth and hence only relatively shallow depths are typically targeted. Fault-zone drilling also allows only discrete sampling of a few select points on a fault surface. Despite these drawbacks, fault-zone drilling has played, and will continue to play, an important role in earthquake science.

Notable deep drilling projects directly related to earthquakes include the SAFOD project, which has drilled through the actively deforming San Andreas Fault zone near Parkfield, California. Among the goals of the project are to drill through the source of a repeating earthquake sequence and to compare it with nearby fault zone materials that exhibit creep rather than earthquakes. Another important fault-zone drilling experiment is the Taiwan Chelungpu fault drilling project. This project has drilled into the shallow part of the rupture of the M_w 7.6 1999 Chi-Chi, Taiwan, earthquake in order to study fault-zone properties in the immediate aftermath of a large earthquake. Other fault-zone drilling projects of note include the Gulf of Corinth Rift Laboratory, which seeks to understand a continental back-arc rifting environment, and the Nankai Trough Seismogenic Zone Experiment, which has the ambitious goal of drilling the accretionary prism of the Nankai Trough in Japan. These and other fault-zone drilling projects will provide important constraints on the fluids, fluid pressure, chemistry, permeability, deformation mechanisms, and stress as they relate to earthquake behavior.

If fault-zone drilling is carried out soon after an earthquake, it is possible to place constraints on the heat produced during faulting, and hence the amount of work done against friction, by measuring a temperature anomaly (Brodsky et al., 2010). This has been carried out for three earthquakes: the Chi-Chi, Taiwan, earthquake (Kano et al., 2006; Tanaka et al., 2006); the 2008 Wenchuan, China, earthquake (Xue et al., 2013); and the 2011 Tohoku-Oki, Japan, earthquake (Fulton et al., personal communication). In each case, the temperature anomaly is small, suggesting that slip occurred at very low levels of frictional resistance. A detailed discussion of fault-zone drilling projects can be found in **Chapter 4.07**.

Another environment that is used to study earthquakes *in situ* is in deep mines. The act of mining removes load-bearing material and induces earthquakes. These earthquakes are extremely dangerous to miners, and hence, there is a great deal of interest in understanding and predicting them. They also have the potential to tell us much about how ordinary tectonic earthquakes work. Seismicity in mines offers some advantages over studying seismicity remotely. The location of mining operations is known and the likely areas of seismicity can be inferred from them. Another advantage is that seismic instruments can be deployed in three-dimensional arrays, which is not possible with strictly surface observations. Seismicity in mines is also at a scale that bridges the laboratory and usual seismological scale of observation. These are the principal motivation behind the multidisciplinary Natural Earthquake Laboratory in South African Mines (NELSAM) project, which is studying earthquakes in the western deep-level South African gold mine.

4.01.5.6 Earthquakes as a Complex System

Earthquakes do not happen in isolation, but interact strongly with each other. The same goes for faults, which might more properly be described as fault systems. A full numerical treatment of fault systems would require realistic geology, fault geometries, rheology, a full range of earthquake sizes, and multiple earthquakes. Although progress is being made on modeling earthquakes in ways that include each of these characteristics, combining them into a single comprehensive model is far beyond current capabilities. There is, however, an approach to studying earthquakes that seeks simplifications to the effect of these properties or characteristics while at the same time attempting to extract the important system-level behavior that affects earthquakes. Studies of this sort treat the behavior of large earthquake populations statistically, with the goal of understanding the macroscopic behavior based on an understanding of the system at a microscopic level. An example of this sort of treatment, which seeks to explain the *b*-value of seismicity distributions as a manifestation of self-organized criticality, is the study of Bak and Tang (1989). A review of this approach to understanding earthquake phenomena can be found in **Chapter 4.24**.

4.01.6 State of Stress, Nucleation, and Triggering

4.01.6.1 State of Stress

The orientation of stress within the Earth can be determined by diverse methods such as earthquake focal mechanisms, borehole breakouts, and the orientation of volcanic dikes. As shown in **Figure 35**, systematic compilations of stress orientations indicate that it is coherent over large scale length and generally consistent with what is expected from ongoing plate tectonic motion (Zoback et al., 1989). The magnitude of these stresses at seismogenic depth, as well as the orientation of stress in the vicinity of major fault zones, is much less clear and the subject of ongoing research.

The lack of a detectable heat flow anomaly centered on the San Andreas Fault (Brune et al., 1969), which is expected under the assumption that fault friction at depth is similar to that

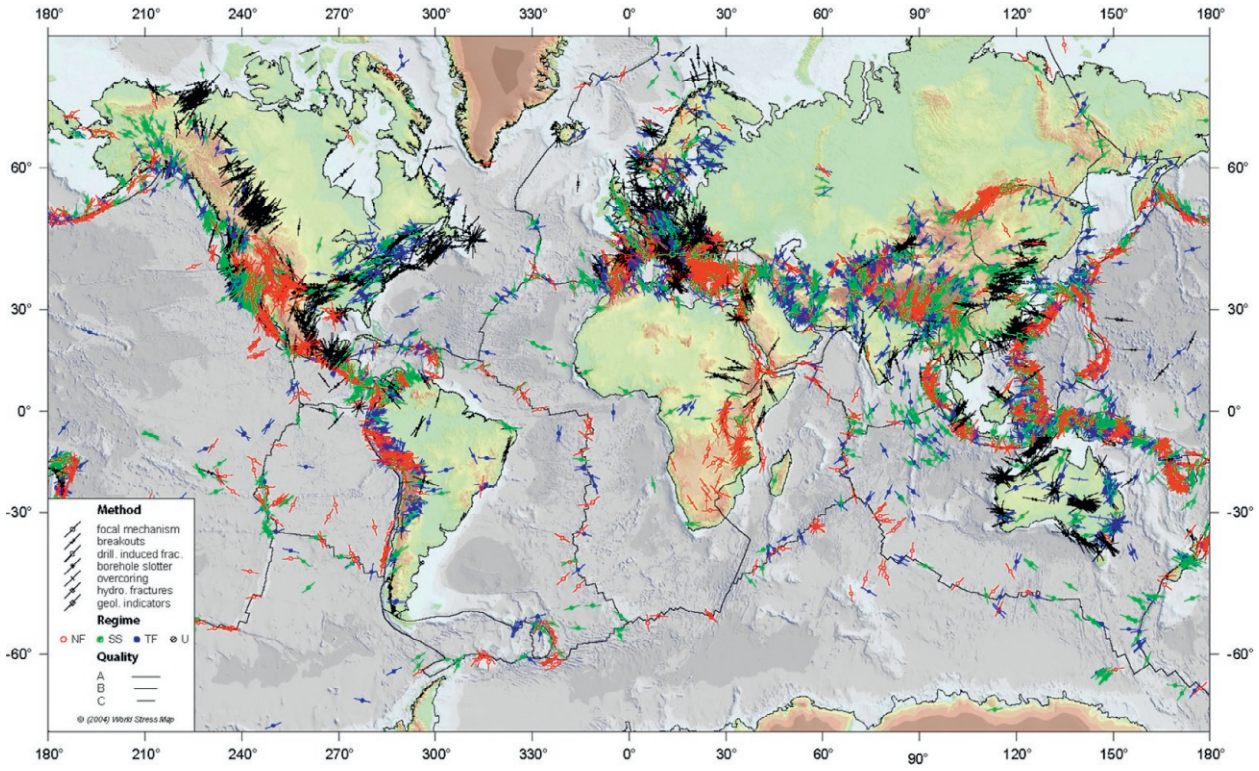


Figure 35 2005 version of the world stress map. Stress directions are determined by a range of techniques listed in the legend. Reproduced from Reinecker J, Heidbach O, Tingay M, Sperner B, and Müller B (2005) The release 2005 of the World Stress Map (available online at www.world-stress-map.org).

observed in the laboratory, indicates that friction is low, at least during slip in large earthquakes. The observation that stress along the San Andreas tends to be oriented such that the maximum compressive stress is close to fault normal (Zoback et al., 1987) is consistent with the notion that fault slip occurs at low levels of shear stress. This conclusion has been disputed by Scholz (2000) who argued that stress rotates near the San Andreas Fault to angles that are consistent with expectations from ordinary levels of friction. Further, borehole stress measurements suggest that stress increases with depth in accord with expectations from the laboratory. Scholz (2000) cited a much broader heat flow anomaly – ~80 km wide – centered on the San Andreas Fault (Lachenbruch and Sass, 1980) as the signature of heat generated during repeated large earthquakes. Larger permeabilities than have been measured in heat flow experiments would be required for convective heat transport to operate using crustal fluids, but if permeability is scale-dependent, then it could reach such an extent. Recent work on geologic signatures of transient heating (d'Alessio et al., 2003) and on transient heating in the immediate aftermath of large earthquakes (Kano et al., 2006) should help to resolve this important and long-standing controversy.

Earthquake focal mechanisms provide constraints on the local orientation of the stress field. Earthquakes can also provide information on changes in stress. During an earthquake, shear stress on the fault changes – for the most part dropping on parts of the fault that slipped and increasing on parts of the fault that did not. If the area of slip can be determined, then a characteristic dimension of faulting can be inferred from it. The ratio of the average slip to the characteristic dimension can be thought of as a strain drop and interpreted through various simple models of source geometry as a drop in stress. This 'stress drop' represents how much the stress decreased from its initial level before the earthquake to the final level after the earthquake – at least on the parts of the fault that slipped. Thus, stress drop provides indirect, model-dependent, but nonetheless useful, constraints on the absolute level of stress. It is usually assumed that friction at depth is sufficiently large that shear stress does not drop below zero and act in the opposite direction following an earthquake, a condition that is termed 'dynamic overshoot.' Under this assumption, stress drops provide lower bounds on the level of shear stress acting on faults at seismogenic depths.

Stress drops (see **Section 4.01.3.4**) are determined seismically from a combination of the seismic moment and some measure of characteristic dimension. For shallow large earthquakes, the characteristic dimension can be estimated from various sources such as geodetic data, geologic data (e.g., surface break), tsunami source, slip inversion, aftershock area, and macroseismic data (e.g., intensity distribution). For small earthquakes, the source dimension cannot be directly determined and is usually taken as the radius of a circular crack. The radius is determined by measuring the corner frequency – the frequency of radiated waves at which the finiteness of the source of the earthquake first becomes apparent – with the assumption that the rupture speed is approximately equal to the shear-wave speed. Data from earthquakes, large and small, show that stress drops follow a scaling relation of the form $M_0 \sim r^3$. Since the faulted

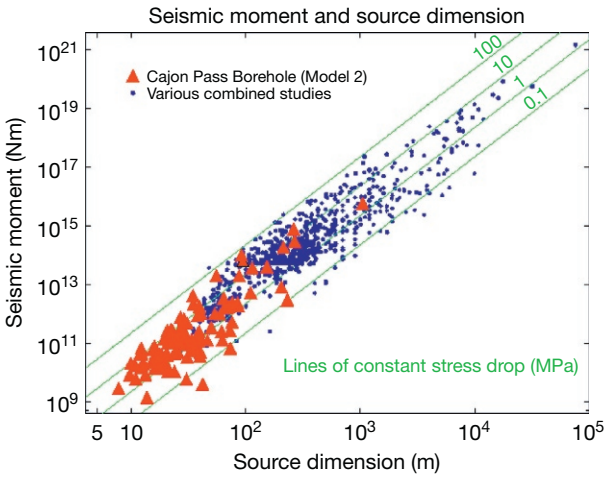


Figure 36 Abercrombie (1995) found from corner frequency measurements taken from a borehole seismometer that the relationship between source dimension and seismic moment for small earthquakes was the same as that for large earthquakes. Diagonal lines showing constant stress drop ($M_0 \sim r^3$) explain the trend in the data.

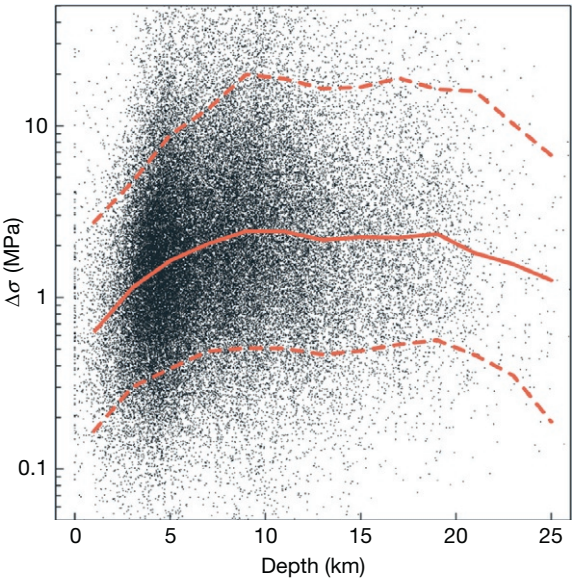


Figure 37 Shearer et al. (2006) were able to discern a depth dependence to stress drop for earthquakes in Southern California. Black dots show individual event stress drops. Solid line shows median values with depth and 10th and 90th percentiles shown as dashed lines.

area, a, follows $a \sim r^2$, this leads to the average slip Δu, scaling as $\Delta u \sim r$, or in other words, the ratio of slip over length is constant (Figure 36).

Assuming that stress drop reflects the ambient stress, it is reasonable to expect that as depth increases, and the stress due to overburden changes with it, earthquake stress drops would increase as well. Such a correlation has proven difficult to find. In part, this may be due to the strong dependence of stress drop estimates on the corner frequency, which means that a small error in corner frequency will lead to a large error in the stress drop. By stacking over thousands of earthquakes, Shearer et al. (2006) had discerned just such a correlation (Figure 37). The depth dependence is much weaker, however, than would be expected if stress drops scaled with the overburden, and most of the dependence is in the upper third of the ~15 km depth extent of the seismogenic zone.

Another approach to determining stress drops is to use repeating earthquakes – earthquakes that repeatedly rupture a particular part of a fault – together with the assumption that these areas slip only seismically and keep up with long-term steady slip. For a given seismic moment, one can determine the average slip such that the slip matches the long-term rate. Stress drops determined by this approach (Nadeau and Johnson, 1998) are strongly scale-dependent. Stress drops for earthquakes of magnitude ~6 are comparable to those based on corner frequency measurements, but they find that stress drops measured using slip-rate matching of repeating earthquakes increase rapidly with decreasing earthquake size such that for the smallest earthquakes, they are one to two orders of magnitude higher.

4.01.6.2 Earthquake Nucleation and Short-Term Earthquake Prediction

Earthquake prediction can fairly be called the holy grail of seismology. While it may be that aspects of earthquakes are inherently unpredictable, the ability to predict behavior is the *sine qua non* of science. Thus, over the long term at least, we can make predictions about earthquakes based on what we really do understand. For example, it is safe to 'predict' that over the next century, most earthquakes will occur on plate boundaries. This is not what most people have in mind when the phrase 'earthquake prediction' is used. Rather, it is understood to mean prediction of the location, size, and time of a specific earthquake to within a matter of days to months. Needless to say, this sort of prediction has not been realized, and there are many who maintain that the nature of the earthquake instability makes prediction on such short timescales intrinsically impossible.

The challenge of short-term earthquake prediction can be illustrated by comparing earthquakes with lightning – a phenomenon that shares some important similarities. Both earthquakes and lightning involve the catastrophic conversion of slowly accumulated potential energy to suddenly released kinetic energy. In the case of lightning, it is a slowly accumulated electrical charge that suddenly flows as electrical current in a lightning bolt and radiates sound waves as thunder. For earthquakes, it is gradually accumulated elastic strain energy stored in the Earth's crust that suddenly accelerates the material on both sides of the fault and radiates energy in the form of seismic waves. The timescale of charge accumulation in a thunderhead can be measured in tens of seconds (order 10^1 s), whereas the timescale of strain accumulation in the Earth's crust can be measured in centuries (order 10^{10} s). Based on the timescale of energy accumulation, predicting the size, location, and time of an earthquake to within a day (order 10^5 s) is equivalent to predicting the amperage, location, and time of a lightning bolt to within a tenth of a millisecond (order 10^{-4} s). At face value, this sounds wildly unrealistic, but with lightning, such a prediction is straightforward because lightning is preceded by a *nucleation* process.

The Earth's lower atmosphere is normally a very effective insulator. In order for a lightning bolt to propagate between a cloud and the ground, it must first create a conductive path for the lightning strike to follow. It does this through the formation of a stepped leader – the piecewise ionization of air that creates a path from cloud to ground. As the air is ionized, it transforms from an insulator to a conductor and forms a channel for the flow of current. Once a channel connects the cloud to the ground, the circuit is completed, charge can flow, and the discharge of the main lightning strike occurs, typically traveling from ground to cloud. For a lightning strike, the formation of the stepped leader is the nucleation phase. The stepped leader immediately precedes the lightning bolt and occurs over the entire length of the subsequent main event. The prediction of the time and location of a lightning bolt would be trivial once the stepped leader is observed. The key question for earthquake prediction is whether or not earthquakes have a similar nucleation process.

Laboratory (Dieterich, 1979a,b) and theoretical (Andrews, 1976) models of earthquake initiation both indicate that aseismic nucleation must occur before slip can propagate unstably across a fault, so it seems highly likely that nucleation of some sort must occur before earthquakes. Assuming it does, the question becomes, how extensive is the earthquake nucleation process? If it occurs over a very limited part of the fault, and thereafter earthquake rupture propagates as a self-sustaining process, then earthquake prediction will be very difficult – perhaps impossible. In this scenario, termed the Cascade model by Ellsworth and Beroza (1995), nucleation of small and large earthquakes might be identical. So for earthquake prediction to become a reality, not only would we need to observe the nucleation process, but also we would have to know that the fault is ready to host a self-sustaining, sometimes termed 'spontaneous,' rupture. This is a great deal more knowledge than we have about faults at present, and such information is likely to be hard to come by. If, on the other hand, the nucleation process scales with the size of the eventual earthquake, for instance, if the size of the nucleation zone were proportional to the size of the eventual earthquake, then earthquake prediction would be a good deal more likely, though still extremely challenging. Another source of significant uncertainty is multiple-fault interaction. The M_w 7.9 Denali, Alaska, earthquake of 2002, for example, began as a reverse faulting event on the Susitna Glacier fault, which led to a large strike-slip fault earthquake on the Denali Fault.

4.01.6.3 Earthquake Forecasting

The prospects for estimating earthquake likelihood over timescales of decades and longer, sometimes known as earthquake forecasting, are considerably brighter. We know that the great majority of earthquakes occur either on plate boundaries or on previously identified fault systems, so for the most part, we know where most earthquakes will occur. We also know quite accurately how rapidly plates are moving, so we have fault slip boundary conditions that must be met over a fault, or a system of faults that comprise a plate boundary, over time. This knowledge, together with information such as the time of the last major earthquake, or assumptions of the size distribution of large earthquakes, is enough to start estimating earthquake likelihood. While earthquake forecasts are made routinely now as a part of seismic hazard assessment, the long timescales over which they apply make it difficult to test them rigorously. Moreover, to the extent that they have been tested, some studies have found them wanting.

The seismic gap hypothesis as formulated by Fedotov (1965), Mogi (1968), McCann et al. (1979), and Kelleher et al. (1973) holds that earthquakes are more likely to occur in areas prone to large earthquakes, but that have not recently had large earthquakes. Kagan and Jackson (1991) found that, as formulated, the seismic gap model has failed to predict where major earthquakes are more likely to occur. They interpreted this to mean that earthquakes are clustered in time, adding that because the theory of plate tectonics only provides long-term constraints on slip rates, it does not imply the regularity of earthquake recurrence.

Another possible interpretation of the failure of the seismic gap hypothesis is that it needs to be refined to take into account complicating factors, such as the possibility that earthquakes are not characteristic or that slip in earthquakes is strongly heterogeneous. As an example, consider the case of the 1981 Playa Azul and subsequent 1985 Michoacan, Mexico, earthquakes (Singh et al., 1988; **Figure 38**). The two earthquakes occurred just 4 years apart and their spatial centroids are essentially the same, yet they clearly had very different, and most probably complementary, slip distributions. It should be possible to test a weaker form of the seismic gap hypothesis, such as whether or not earthquakes occur in areas that have not slipped recently, once geologic fault slip rates are taken into account. The Playa Azul and Michoacan earthquakes would pass this test while failing the original seismic gap hypothesis. It is interesting to note that the other seismic gap shown in **Figure 38**, in the state of Guerrero, has not been the source of a large earthquake in the interval since Singh et al. (1988) published their paper. The Guerrero gap, however, has been the site of a $M_w \sim 7.5$ large silent earthquake (Kostoglodov et al., 2003), an observation that complicates the interpretation of this seismic gap. With the advent of GPS, earthquake forecasting has become significantly more quantitative.

A good example of forecasting is for the 2010 $M_w = 8.8$ Maule, Chile, earthquake. The earthquake occurred on the Chilean megathrust where another large earthquake had occurred in 1835 ($M \approx 8$–8.5). Charles Darwin experienced this earthquake and wrote a firsthand account of it in his 1845 book. Since that time, no major earthquake had occurred on this segment, called the Charles Darwin gap. The average repeat time of Chilean megathrust earthquakes is estimated to be 100–200 years. Thus, in the 1990s, the Charles Darwin gap had been expected to host another large earthquake in the near future. These are the boundary and the initial conditions to be used for forecasting the seismic activity in this gap. Having realized these conditions, several investigators began taking GPS measurements in this general area in 1996. Based on the displacement field determined by the 10-year GPS survey, Ruegg et al. (2009) concluded that if the plate coupling has been 100% since 1835, their results indicated ~ 10 m of slip deficit along this gap (slip deficit is the amount of slip that would occur if the accumulated strain were released in an earthquake) and forecasted that a M_w 8.5 earthquake would occur if the gap ruptured soon. Then, on 27 February 2010, a

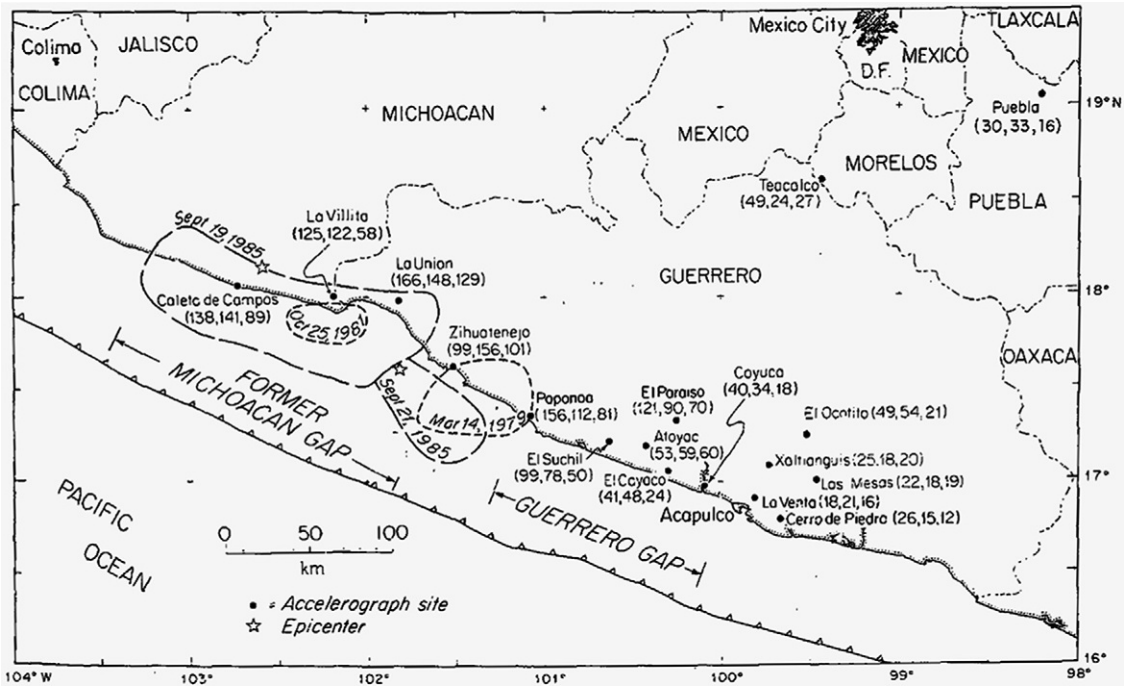


Figure 38 Map showing aftershock zones of the September 19 and 21 Michoacan, Mexico, earthquakes. Note that the rupture of the two earthquakes abuts, but that the 1985 rupture is much larger than, and entirely envelopes, the rupture zone of the 25 October 1981 Playa Azul earthquake. If such behavior is common, then it will greatly complicate attempts to apply the seismic gap hypothesis. Reproduced from Singh SK, Mena E, and Castro R (1988) Some aspects of source characteristics of the 19 September 1985 Michoacan earthquake and ground motion amplification in and near Mexico City from strong motion data. *Bulletin of Seismological Society of America* 78: 451–477.

$M_w = 8.8$ earthquake did occur in the expected area with the expected mechanism (e.g., Delouis et al., 2010; Lay et al., 2010; Lorito et al., 2011; Vigny et al., 2011). The magnitude was slightly larger than what had been forecast, but some uncertainty is inevitable considering the complexities of the subduction zone structure and the stochastic elements associated with fracture phenomena like an earthquake.

In contrast to the case of the 2010 Chile earthquake, the case of the 2011 $M_w = 9.0$ Tohoku-Oki earthquake illustrates the difficulty in forecasting of very complex natural processes like earthquakes even if extensive GPS data are available. For the subduction zone off the Tohoku coast, many GPS studies had previously been conducted in which onshore GPS data were extensively modeled (e.g., Hashimoto et al., 2009; Ito et al., 2000; Loveless and Meade, 2010; Mazzotti et al., 2000; Nishimura et al., 2004; Suwa et al., 2006). All results indicated evidence for large strain accumulation offshore. However, because of the limited spatial coverage (i.e., only onshore) of strain measurements, the location of strain accumulation could not be determined very well, especially in the east–west direction. Furthermore, because of the presence of many $M \sim 7.5$ earthquakes in the last several hundred years of historical record off Miyagi Prefecture, and the absence of direct evidence for, or possibly irregular occurrence of, extremely large earthquakes, together with the spatially heterogeneous mechanical or frictional properties of the plate boundary, the seismic potential of this region had been poorly understood, and the devastating $M_w = 9.0$ earthquake occurred as a big surprise to most seismologists. Thus, this example clearly demonstrates the difficulty in making accurate scientific forecasts of great earthquakes even with extensive GPS data. Nevertheless, as we improve our knowledge by ocean-bottom measurements and by studying many earthquakes in other tectonic environments, we will be able to reduce these uncertainties.

4.01.6.4 Static Stress Triggering

Over the past few decades, it has become increasingly obvious that our ability to forecast earthquakes is strongly dependent on our understanding of how earthquakes interact with one another. This is a critical question for earthquake hazard assessment that will inevitably depend on essential aspects of earthquake physics. How does the occurrence of one earthquake influence the occurrence of another? The seismic gap hypothesis is a limiting case of this in which the effect of an earthquake on a subsequent earthquake at the same place on the plate boundary is considered. Beyond this, the most obvious such interaction is an aftershock sequence in which many smaller earthquakes occur in the aftermath of a larger 'mainshock.' Large earthquakes are sometimes implicated in the triggering of other large earthquakes too, with the most famous example being the twentieth-century sequence of large earthquakes that eventually ruptured most of the North Anatolian Fault in Turkey (Toksöz et al., 1979).

Earthquake interaction is thought to occur through stress triggering, that is, failure in a future earthquake is either encouraged or discouraged by the static stress change induced by an earthquake that precedes it. Invoking the same criterion as in the subsection on faulting and pore fluids, fault slip will occur when the shear stress, τ, acting across the fault exceeds

the sum of the cohesion, C, and the effective normal stress multiplied by the coefficient of friction, μ:

$$\tau > \mu(\sigma - P) + C$$

where the term in the parentheses is referred to as the effective normal stress and is the difference between the fault-normal stress, σ, and the pore pressure, P. The cohesion is usually assumed to be negligible. By this criterion, earthquakes can be triggered by either an increase in shear stress, τ; a decrease in normal stress, σ; an increase in the pore pressure, P; or some combination of the three. There is now a substantial body of literature that explores static stress triggering (Stein, 1999) and its implications for earthquake physics, for earthquake interaction, and for time-dependent seismic hazard assessment. **Figure 39** shows perhaps the clearest example of static stress triggering: the progression of earthquakes across the North Anatolian Fault in Turkey during the twentieth century (Stein et al., 1997). As outlined in **Section 4.01.5.1**, friction on faults is thought to be dependent on slip rate and the past slip history of a fault. Changes in long-term earthquake probabilities due to stress steps induced by nearby earthquakes under rate- and state-variable friction laws have been explored by Parsons et al. (2000).

Perhaps, less obvious than the triggering of earthquakes by static stress changes is the notion of a stress shadow, in which earthquakes are inhibited because the stress necessary to drive them is relieved by a nearby large earthquake. The most famous and convincing example of this may be the almost complete lack of significant earthquakes in Northern California following the 1906 San Francisco earthquake (Ellsworth et al., 1981). Over the 100 years since the 1906 earthquake, there have been several earthquakes near the ends of the 1906 earthquake rupture – most notably the 1989 Loma Prieta earthquake, but nowhere near the level of seismicity that occurred in the 50 years leading up to the 1906 earthquake (Working Group and on California Earthquake Probabilities, 2002). Stress shadows are important because they provide an opportunity to distinguish between static triggering and dynamic triggering (next section). The presence of stress shadows is strong supporting evidence for the importance of static stress changes in triggering earthquake activity (Toda et al., 2012). The topic of static stress change triggering of earthquakes is covered in detail in **Chapter 4.10**.

4.01.6.5 Dynamic Triggering

The examples cited earlier are of earthquakes either encouraging or inhibiting other earthquakes as a consequence of the static stress change that they induce. Another mode of earthquake interaction is through dynamic stress changes, that is, the transient stresses that are transmitted in the form of seismic waves. These were first clearly recognized in the aftermath of the 1992 Landers, California, earthquake in which 'aftershocks' were triggered at extreme distances of more than 1000 km from the mainshock (Hill et al., 1993). At these distances, static stress changes induced by the earthquake are much smaller than the periodic stress changes induced by solid Earth tides and are thus unlikely to act effectively as a trigger. Dynamic stress changes, on the other hand, are still quite substantial at great distances. This is attributable to the fact that static stress changes decay approximately as distance cubed once a distance comparable to the source extent is exceeded. In contrast, dynamic stresses associated with surface waves decay much more slowly.

Since the Landers earthquake, dynamic triggering at great distance has been documented in the 1999 Izmit earthquake, the 2002 Denali earthquake, the 2004 Sumatra earthquake, and other events. In the case of Sumatra, microearthquakes were triggered at Mt. Wrangell, Alaska, at a distance of nearly 11 000 km from the rupture zone. In this instance, the triggered earthquakes were observed to occur during the maximum vertical displacement of the large and extremely long-period Rayleigh waves (**Figure 40**) generated by that event (West et al., 2005).

4.01.6.6 Temporal Distribution of Earthquakes

Large earthquakes are typically followed by numerous aftershocks. The rate of aftershocks R, as a function of time, t, following a large earthquake typically decays in time following a relation known as Omori's law, which in its modified form (Utsu, 1961), can be written

$$R(t) \sim (t + c)^{-p}$$

where c is a constant and the exponent, p, is usually observed to be $\sim$1. Aftershocks are usually thought to occur in response to the stress change imposed by the mainshock they follow (Scholz, 1990). The mechanism behind the temporal decay of Omori's law has been attributed to many physical effects. Elastic effects act immediately and do not explain the gradual decay. Among the mechanisms that have been proposed to explain the temporal decay of aftershock rate are pore fluid flow (Nur and Booker, 1972), viscoelastic relaxation (Freed and Lin, 2001), stress corrosion (Scholz, 1968), and earthquake nucleation under rate- and state-variable friction (Dieterich, 1994).

Since dynamic stresses can trigger earthquakes at large distance, it stands to reason that they can trigger earthquakes at short distances as well. This might explain the diverse aftershocks of the 1989 Loma Prieta earthquake and the observation that they were not favored by the mainshock-induced stress change (Beroza and Zoback, 1993). Felzer and Brodsky (2006) showed that the distance decay of aftershocks was consistent with dynamic triggering, even for small, local earthquakes.

Whether by static triggering, dynamic triggering, or other processes such as poroelastic effects, it is clear that earthquakes interact with each other. That is, the occurrence of one earthquake affects the probability of other earthquakes. This is true not only for large earthquakes but also for small earthquakes. Models that attempt to document and quantify the consequences of these interactions have come to be known by the acronym 'ETAS,' which stands for epidemic-type aftershock sequences (Ogata, 1988). These and other seismicity models are covered in more detail in **Chapter 4.24**.

4.01.7 Associated Problems

4.01.7.1 Strong Motion Prediction

From a societal perspective, the prediction of strong ground motion is arguably one of the most important issues that

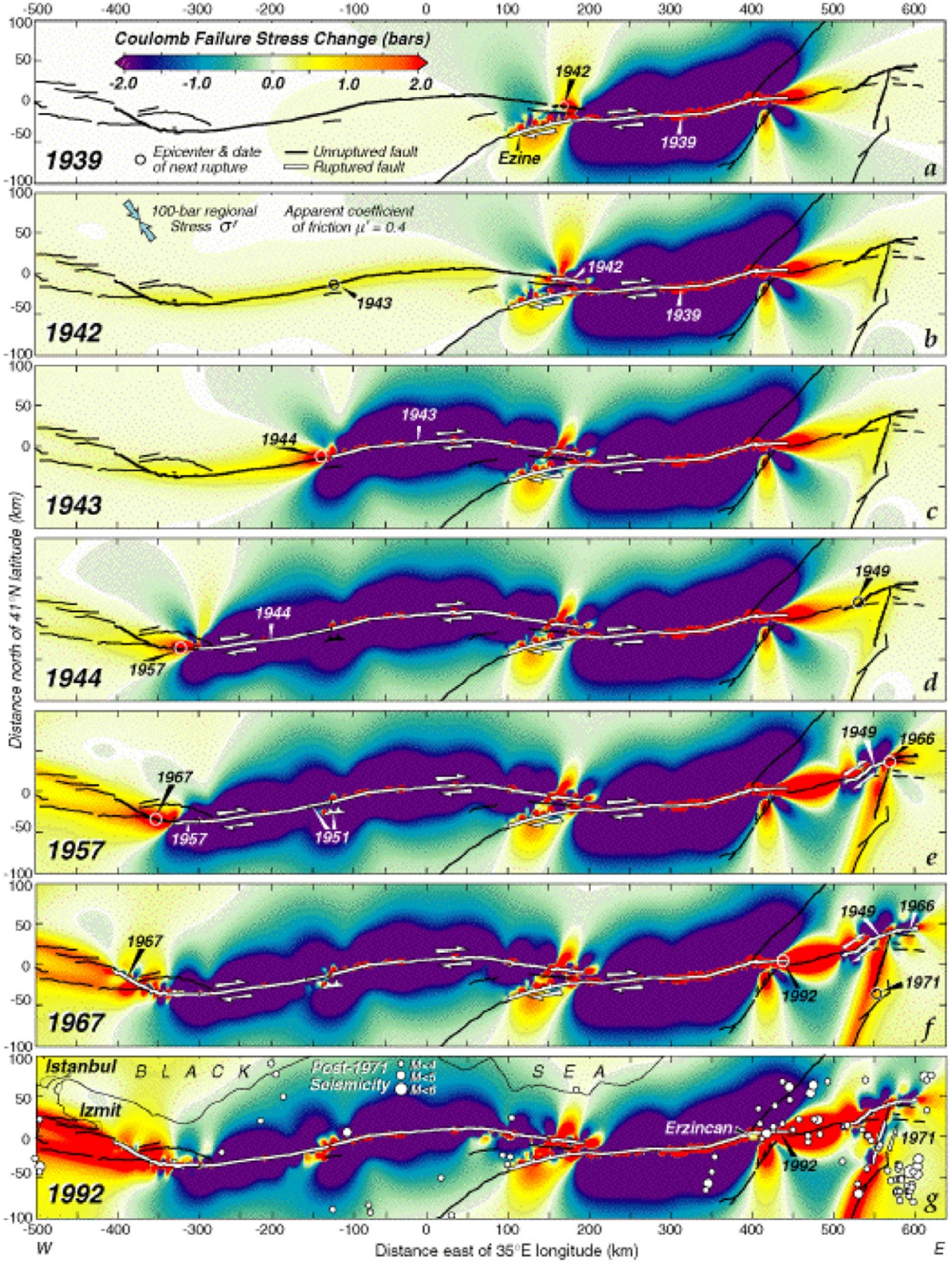


Figure 4 17 Oct 96 **Stein et al.**

Figure 39 Cumulative Coulomb stress changes caused by large earthquakes on the North Anatolian Fault in Turkey since 1939. In each panel, the epicenter of the next earthquake to rupture is circled. All but the 1943 epicenter lie in regions where the stress rose significantly, typically by 2–5 bar, owing to a combination of previous earthquakes and deep fault slip. Relocated earthquakes (Engdahl et al., 1998) are shown in the last panel, with the 1992 aftershock sequence removed. Calculations are made assuming a Poisson, elastic half-space. Slip is assumed uniform from 0 to 12.5 km depth. Fault is projected such that regional stress maintains a nearly fixed angle relative to the fault. Reproduced from http://quake.usgs.gov/research/deformation/modeling/papers/anatolia.html.

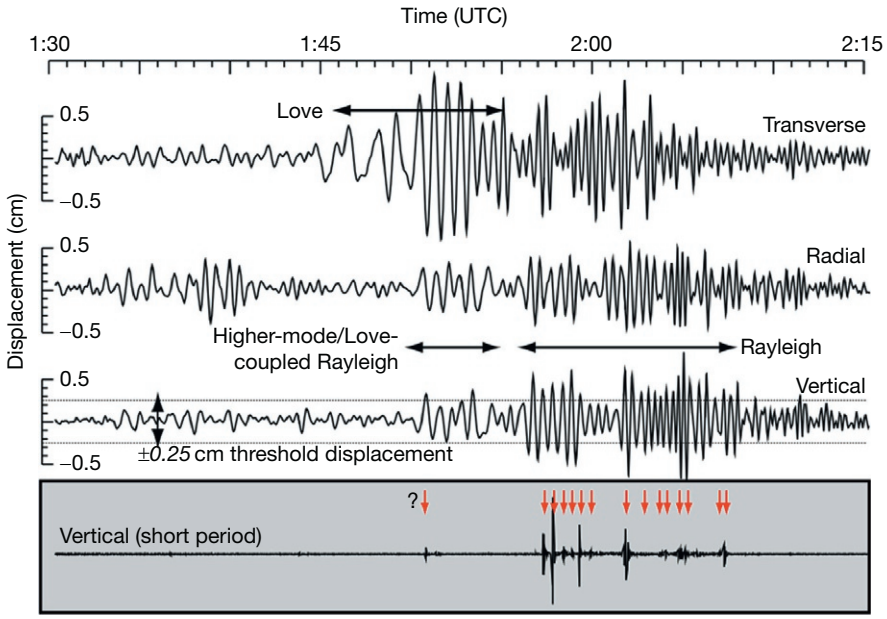


Figure 40 Three-component displacement records for the Sumatra earthquake as recorded at PAX. The bottom trace is from station WANC and has been filtered on 0.5–20 Hz to show timing of local events at Wrangell Volcano. The timing and spacing of the triggered events (highlighted with red arrows) with the Rayleigh wave phase and amplitude indicate Rayleigh motion is responsible for the triggering. Despite comparable amplitudes, there is no such correlation for the Love wave (screen grab from http://www.aeic.alaska.edu/input/west/wa_trigger/).

seismologists can address. Much of the risk that earthquakes pose, particularly to property, could not be mitigated by short-term prediction, which is an area of seismological research that is afforded a lot more attention by the general public. Even if short-term earthquake prediction should someday prove possible and reliable, it would not be possible, for example, to retrofit large engineered structures on a timescale of days to weeks.

The wave propagation properties of the Earth's crust are strongly heterogeneous on all scales. As a result, the seismic wave field becomes strongly distorted as it propagates through the crust. Compounding this wave field complexity is the possible complexity of the earthquake source itself, which can be substantial for the large damaging earthquakes that are of concern in strong ground motion prediction. Understanding and predicting this variability in strong ground motion pose a major challenge to engineering seismologists and civil engineers who have to design structures based on anticipated ground motions.

So how is strong ground motion predicted? Work in this area bridges the disciplines of seismology and earthquake engineering and thus is termed 'engineering seismology.' Engineering seismology seeks to predict the intensity of ground shaking likely to occur in future earthquakes. The standard procedure for predicting strong ground motion proceeds as follows.

The first step is to decide what intensity measure is to be used, which is an engineering decision based on what information is needed for structural design. Commonly used intensity measures include PGA, peak ground velocity, spectral acceleration (Sa), spectral velocity (Sv), and spectral displacement (Sd). Then, a dataset of strong motion recordings is assembled. A great deal of care is taken in the process to ensure that the ground motion parameter of interest is not somehow obscured by particular characteristics of the recording site. Moreover, reliable information such as the distance to the fault plane and the geologic characteristics of the recording site are required in order to take full advantage of the data.

Once these data are assembled, and the intensity measure is determined for them, we posit a parametric relationship between relevant aspects of the earthquake of interest and the strong ground motion intensity measure. These relationships are usually called 'attenuation relations' despite the fact that they are dominated by effects other than anelastic attenuation. Examples of the factors that are used to develop these relationships include magnitude, closest distance to fault rupture, earthquake mechanism, and the nature of the soils at the site of interest. Next, a regression between these factors and the intensities observed in earthquakes is carried out in order to determine the coefficients of the attenuation relation.

The current scarcity of strong motion data at short distances from large earthquakes means that there are inadequate data to represent the hazard from the most dangerous events. Ground motion computer simulation provides a way to fill this gap in the data. Another argument in favor of ground motion simulation stems from the need of civil engineers for a more complete description of ground motion (i.e., seismograms) instead of a scalar intensity measurement, such as the spectral acceleration.

Simulation of near-fault ground motion that is accurate enough to be used to take the place of attenuation relations is an emerging technology that is not quite ready for application. A major source of uncertainty for these efforts remains an accurate characterization of the earthquake source and, in particular, the relationship between source parameters implied by rupture dynamics (Guatteri et al., 2004). Pseudodynamic

models are useful for incorporating the effects of rupture physics in attempts to simulate strong ground motion (Graves et al., 2011). Also crucial to strong ground motion prediction is the ability to model path effects, that is, the effect of wave propagation through the complex geologic structure of the Earth's crust. This includes not only elastic wave propagation but also loss of wave energy due to anelasticity and changes to the wave field due to nonlinear site effects. The approach of using small earthquakes as empirical Green's functions (eGfs) is a very promising avenue of research for the prediction of high-frequency strong ground motion (Joyner and Boore, 1986). The advantage of the eGf approach is that it includes the true complexity of wave propagation effects. In practice, its application is limited by the availability of suitable eGf events and the possibility that nonlinearities present during the strong ground motion of the mainshock will not affect the relatively weak motions of the eGf event.

4.01.7.2 Tsunamis

A tsunami is a long-period wave train that propagates as a shallow water wave in the open ocean. By 'shallow,' we mean that the wavelength is much longer than the depth of the ocean, such that particle trajectories in the wave are nearly horizontal. This criterion is easily satisfied by earthquake-generated tsunamis, which can have wavelengths 100 s of kilometers long – an order of magnitude greater than the depth of even the deepest ocean trenches. Because they are shallow water waves, tsunamis propagate at high speed across the open ocean – at speeds that approach the speed of travel by commercial jets.

Earthquakes represent one of the primary triggers of tsunamis, but tsunamis can be generated in many ways. The permanent deformation of the sea floor due to an earthquake is the source of some of the largest, most destructive tsunamis, but other important potential tsunami sources include submarine landslides, volcanic eruptions, and bolide impacts. For some earthquakes, such as the 1946 earthquake in the Aleutian Islands, tsunami modeling suggests that landslides triggered by earthquakes may have been the principal tsunami source (Fryer et al., 2004), though Lopez and Okal (2006) modeled the tsunami generation using a very large, but slow earthquake rupture. Clearly both triggered landslides and deformation of the sea floor due to earthquake rupture must be considered capable of tsunami generation regardless of whether one particular tsunami arose from one source or the other. For a detailed treatment of tsunamis as they relate to earthquakes, see **Chapter 4.19**.

4.01.7.3 Test Ban Treaty Verification

Since a nuclear bomb was first tested underground in 1957, seismology has played a key role in nuclear monitoring and Test Ban Treaty verification. Much of the impetus and funding to establish the WWSSN in the early 1960s and its successor the Global Seismic Network nearly three decades later came not from the need to monitor earthquakes, but from the need to monitor nuclear testing. The proliferation of nuclear weapons to additional countries, and the reduction in the target monitoring threshold from 150 kt of TNT to 10 kt, required a truly global seismological network that spans all the continents and records data over a broad range of frequencies. The multilateral nature of the Comprehensive Test Ban Treaty also meant that the data had to be openly available to all interested parties. This combination has served the purposes of Test Ban Treaty verification research, but it has also been a bonanza to earthquake seismology in general. It is also fair to say that earthquake monitoring has been an important source of sustenance for these same networks.

There are a number of issues that arise in Test Ban Treaty verification research. First, and perhaps foremost, is the need to detect and locate small magnitude events that represent possible low-yield nuclear tests. The second task is to discriminate explosions from earthquakes (and, in addition, nuclear tests from other explosions). **Figure 35** shows a comparison of a seismogram from an earthquake with one from an explosion. The relatively efficient generation of P-waves relative to S-waves is one means for discriminating explosions from earthquakes. A third task is to assess the feasibility of various evasion scenarios in which a clandestine test could be sufficiently decoupled from the Earth or somehow disguised such that it would go unreported. As the size threshold decreases, all of these tasks become more difficult due to reduced signal-to-noise ratio, a larger number of small earthquakes, and an increasing population of ordinary chemical explosions in the same size range. Yet another task is to estimate the explosive yield of a test. This was a critical issue for the Threshold Test Ban Treaty in the 1970s and 1980, but it continues to be important. Through precise relative amplitude measurements, the yield of the 2013 test in North Korea has been found to be 12 kt of TNT energy equivalent (Zhang and Wen, 2013). This follows the 2006 test, which is widely regarded as a fizzle, and it is much higher than the 2009 North Korea test, which was only slightly larger than 2 kt (Zhao et al., 2012) (**Figure 41**).

4.01.7.4 Solid Earth–Atmospheric Coupling

Seismic waves are generated not only by earthquakes but also by other sources. Volcanoes, for example, are obviously prodigious sources of earthquakes and other seismic sources such as harmonic tremor. Large landslides on volcanoes are another source. All of these represent processes within the solid Earth. Vigorous volcanic eruptions disturb the atmosphere sufficiently that the acoustic and gravitational waves in the atmosphere couple to the solid Earth and excite seismic waves (Kanamori and Mori, 1992; Widmer and Zürn, 1992). The 15 February 2013 Chelyabinsk meteorite was the largest since the 1908 Tunguska event. The airburst caused damage and injuries. It also generated Rayleigh waves detectable at distances of up to 4000 km that were equivalent in amplitude to those from an M_S 3.7 earthquake (Tauzin et al., 2013).

The impedance of the atmosphere is much smaller than that of the solid Earth. As a result, a displacement of the Earth's surface will translate more or less directly into a displacement of the atmosphere. This effect is so small that it is usually neglected, but there are circumstances in which it is important and has a detectable effect. The density of the atmosphere decreases exponentially with increasing altitude. As a result, a perturbation that is modest at the Earth's surface grows dramatically as it propagates upward through the atmosphere. For

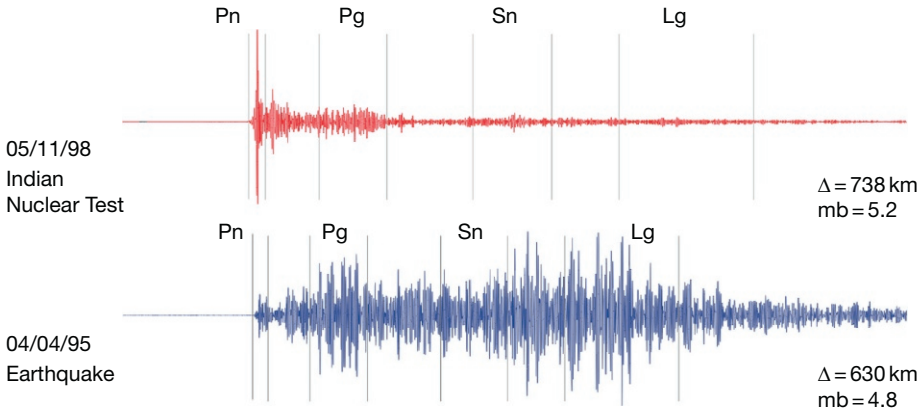


Figure 41 Seismograms recorded on the same instrument at Nilore, Pakistan. Upper seismogram shows ground motion due to an Indian nuclear test. Lower seismogram shows seismogram due to an earthquake at a similar distance. The phases Pn, Pg, Sn, and Lg are shown with vertical lines. Note that the S-wave arrivals from the earthquake are systematically much larger than for the explosion. Clues such as these help to discriminate underground nuclear tests from earthquakes. Reproduced from Walter WR, Rodgers AJ, Mayeda K, Myers SC, Pasyanos M, and Denny M (1998) Preliminary regional seismic analysis of nuclear explosions and earthquakes in Southwest Asia. In: *Proceedings of the 20th Annual Seismic Research Symposium on Monitoring a Comprehensive Test Ban Treaty*. Department of Defense and Department of Energy Report, pp. 442–452.

a major earthquake, this leads to a large and easily detectable change in the ionosphere. Because the ionosphere has a strong effect on electromagnetic waves propagating through it, changes in the ionosphere caused by earthquakes can be detected using, for example, transmissions from GPS satellites (Calais and Minster, 1995). The displacements caused by a large tsunami have also generated detectable ionospheric perturbations (Artru et al., 2005).

The continuous excitation of the very long-period free oscillations (Suda et al., 1998) also appears to result from coupling between the Earth and its atmosphere and/or oceans. The basic observation is that the Earth's very low-frequency fundamental spheroidal modes of oscillations are excited even during periods where there are no earthquakes to account for them. The excitation has an annual signature that strongly suggests it is not driven by processes within the solid Earth. Early explanations for this excitation focused on atmospheric disturbances as the cause (Kobayashi and Nishida, 1998), but more recent analysis suggests excitation by oceanic infragravity waves (**Figure 42**) is more likely (Rhie and Romanowicz, 2004). For more details on this subject, see **Chapter 4.16**.

4.01.8 Earthquake Risk Mitigation

Efficient mitigation of the risk posed by earthquakes requires as a foundation a quantitative measurement of the hazard that earthquakes pose. Once this is understood, informed policy decisions on building codes and their enforcement and on priorities for retrofitting susceptible structures can be made. Much of the work in earthquake hazard analysis casts strong ground motion prediction as a probabilistic statement in a process called probabilistic seismic hazard analysis (PSHA).

PSHA was developed to integrate multiple factors that control the likelihood and strength of shaking to be expected, each of which can be treated as a stochastic process. PSHA asks, for example, what is the ground motion that has a 2% probability of being exceeded over a 50-year time interval? To determine this number, and to map it spatially, requires combining information on earthquake likelihood for all relevant earthquake sources (i.e., active nearby faults) with strong motion attenuation laws that include site effects. The output of PSHA is a map of ground motion intensities at a specified exceedance probability that can be used in building codes to develop design criteria for buildings and by policymakers to develop priorities for earthquake risk reduction. **Figure 43** shows an example calculation for the entire planet.

Given the importance of PSHA to earthquake risk characterization and mitigation, it is important that it be tested to the extent possible. This is challenging because the timescales involved are long, and the probabilities are low. Under the ergodic assumption, it is possible to supplement the short time record of recorded ground motions, by considering a large spatial area. Tests of this kind consider PSHA predictions against observed ground motion (Stirling and Petersen, 2006) or extend it into the preinstrumental era using historical catalogs and ground motion prediction equations (Ward, 1995).

Another important area in seismic risk reduction is earthquake early warning. Earthquake early warning systems rely on the fact that although seismic waves travel at high speed, information that an earthquake has occurred can be transmitted at the speed of light, such that an alert that a significant earthquake is under way can be issued before damaging seismic waves arrive. Earthquake early warning systems are operational in Japan, Mexico, and Taiwan and are in the planning stages in Romania, Turkey, and the United States. The amount of warning that such systems can provide for large earthquakes under favorable circumstances can be several 10 s of seconds. For more details on seismic hazard and real-time warning systems, see **Chapter 4.23**.

4.01.9 Conclusions

In this chapter, we have taken a synoptic view of earthquake seismology and several closely related fields of study. In the

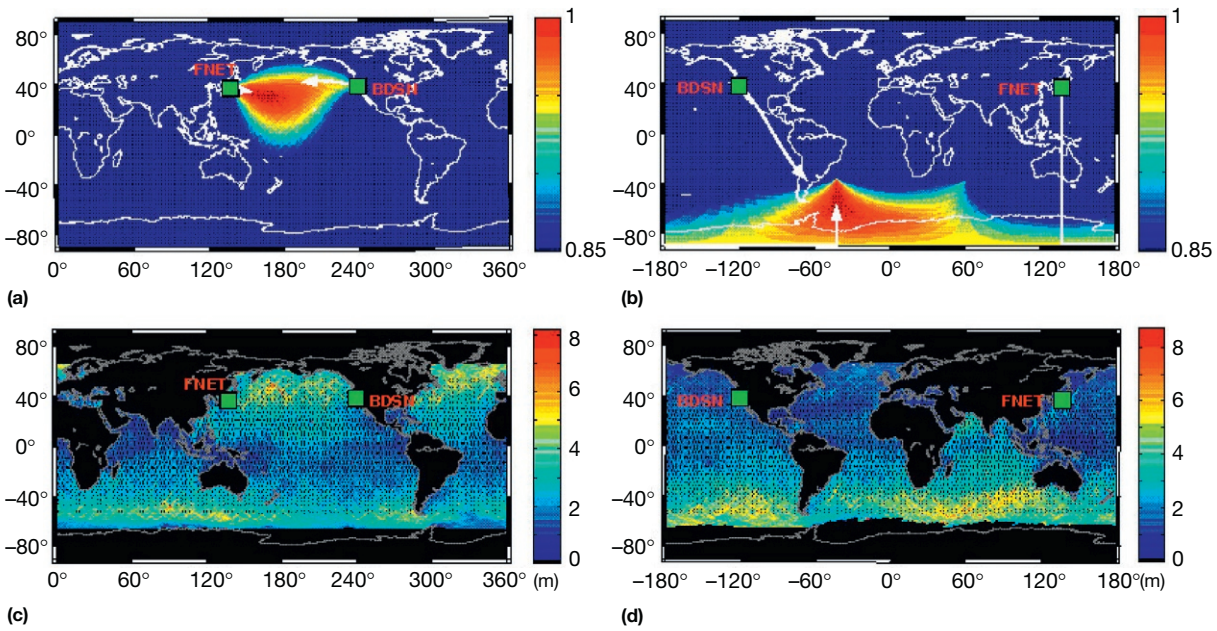


Figure 42 Localization of the source of background excitation of the Earth's free oscillations (Rhie and Romanowicz, 2004) using arrays in Japan and the United States. Arrows in upper panels point from the arrays to the source of excitation (red area) for two different time periods. The excitation on the left is for the northern Pacific and on the right for the southern Atlantic. Sources are more active in boreal and austral winter seasons. Lower panels show monthlong significant wave heights for those seasons from TOPEX/Poseidon measurements.

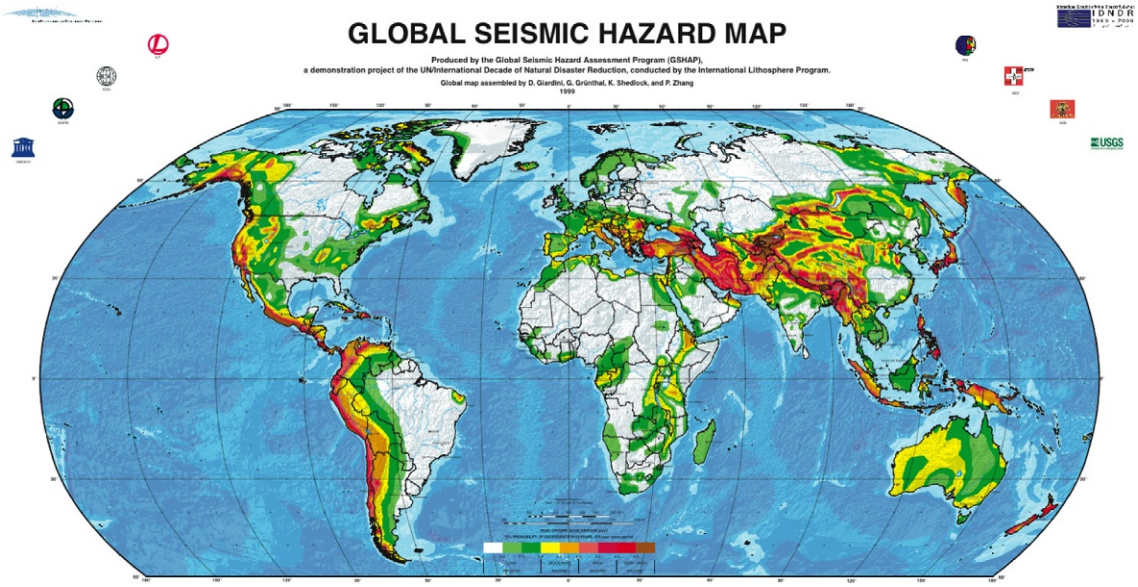

GLOBAL SEISMIC HAZARD MAP
Produced by the Global Seismic Hazard Assessment Program (GSHAP),
a demonstration project of the UN/International Decade of Natural Disaster Reduction, conducted by the International Lithosphere Program.
Global map assembled by D. Giardini, G. Grünthal, K. Shedlock, and P. Zhang
1999
USGS

Figure 43 Global seismic hazard map. Areas of red are of highest seismic hazard, while clear areas are of lowest seismic hazard. Seismic hazard is dominated by plate boundary seismicity, except in continental collisions zones – most notably the Alpide belt. Reproduced from the home page of the global seismic hazards assessment program, http://www.seismo.ethz.ch/GSHAP/.

chapters that follow, many of the same topics are covered, but in much greater detail. The field of earthquake seismology is rapidly evolving. This evolution is driven by several factors. Perhaps, the most important single factor is rapid progress in the quality, density, and wide availability of seismic data. Seismic data centers, which house many terabytes of seismic data and have easy availability via the Internet, have resulted in access for interested seismologists to literally millions of seismograms. These large data volumes enable fundamentally new ways of looking at earthquakes. They also mean that visual analysis of seismograms, except on a proof-of-concept basis, is rapidly becoming a thing of the past. The computational

capabilities of computers have kept pace with the rapid rise in data availability. As data collection and computational resources continue a rapid upward trajectory, imaginative seismologists are presented with an unprecedented opportunity to pose, and to answer, new questions about the earthquake process.

As a result, we can look forward to rapid progress in the field of earthquake and volcano seismology. This progress will be punctuated by large, devastating earthquakes. As in the past, important aspects of these earthquakes will not be expected, which testifies to the fact that earthquake science remains a young and exciting field of research. The havoc future earthquakes will wreak and will also serve to remind seismologists, and those who depend on their research, of the profound societal relevance of earthquake science.

References

Aagaard BT and Heaton TH (2004) Near-source ground motions from simulations of sustained intersonic and supersonic fault ruptures. *Bulletin of the Seismological Society of America* 94: 2064–2078.

Abercrombie RE (1995) Earthquake source scaling relationships from −1 to 5 ML, using seismograms recorded at 2.5 km depth. *Journal of Geophysical Research* 100: 24015–24036.

Abercrombie RE, McGarr A, Kanamori H, and Di Toro G (eds.) (2006) *Earthquakes: Radiated Energy and the Physics of Faulting*, In: *AGU Geophysical Monograph Series* 170: p. 327.

Abercrombie RE and Rice JR (2005) Can observations of earthquake scaling constrain slip weakening? *Geophysical Journal International* 162: 406–424. http://dx.doi.org/10.1111/j.1365-246X.2005.02579.x.

Aki K (1966) Generation and propagation of G waves from the Niigata earthquake of June 16, 1964: Part 2. Estimation of earthquake moment, released energy and stress drop from the G wave spectra. *Bulletin of the Earthquake Research Institute, University of Tokyo* 44: 73–88.

Andrews DJ (1976) Rupture velocity of plane strain shear cracks. *Journal of Geophysical Research* 81: 5679–5687.

Andrews DJ (2005) Rupture dynamics with energy loss outside the slip zone. *Journal of Geophysical Research* 110. http://dx.doi.org/10.1029/2004JB003191.

Andrews DJ and Ben-Zion Y (1997) Wrinkle-like slip pulse on a fault between different materials. *Journal of Geophysical Research* 102: 553–571.

Argus DF and Gordon RG (2001) Present tectonic motion across the Coast Ranges and San Andreas fault system in central California. *Geological Society of America Bulletin* 113: 1580–1592.

Artru J, Ducic V, Kanamori H, Lognonne P, and Murakami M (2005) Ionospheric gravity waves induced by tsunamis. *Geophysical Journal International* 160: 840–848. http://dx.doi.org/10.1111/j.1365-246X.2005.02552.x.

Arvidsson R (1996) Fennoscandian earthquakes: Whole crustal rupturing related to postglacial rebound. *Science* 274: 744–746.

Astiz L, Lay T, and Kanamori H (1988) Large intermediate-depth earthquakes and the subduction process. *Physics of the Earth and Planetary Interiors* 53: 80–166. http://dx.doi.org/10.1016/0031-9201(88)90138-0.

Backus G and Mulcahy M (1976) Moment tensors and other phenomenological descriptions of seismic sources II. Discontinuous displacements. *Geophysical Journal of the Royal Astronomical Society* 47: 301–329.

Bak P and Tang C (1989) Earthquakes as a self-organized critical phenomenon. *Journal of Geophysical Research* 94: 15635–15637.

Bakun WH and Wentworth CM (1997) Estimating earthquake location and magnitude from intensity data. *Bulletin of Seismological Society of America* 877: 1502–1521.

Barazangi M and Dorman J (1969) World seismicity maps compiled from ESSA, Coast and Geodetic Survey, epicenter data, 1961–1967. *Bulletin of Seismological Society of America* 59: 369–380.

Beck ME (1983) On the mechanism of tectonic transport in zones of oblique subduction. *Tectonophysics* 93: 1–11.

Ben-Zion Y, Peng Z, Okaya D, et al. (2003) A shallow fault-zone structure illuminated by trapped waves in the Karadere-Duzce branch of the North Anatolian Fault, Turkey. *Geophysical Journal International* 152: 699–717.

Beroza GC and Zoback MD (1993) Mechanism diversity of the Loma Prieta aftershocks and the mechanics of mainshock-aftershock interaction. *Science* 259: 210–213.

Bilham R (1995) Global fatalities in the past 2000 years: Prognosis for the next 30. In: Rundle J, Klein F, and Turcotte D (eds.) *Reduction and Predictability of Natural Disasters. Santa Fe Institute Studies in the Sciences of Complexity,* 25, pp. 19–31. Boston, MA: Addison-Wesley.

Boatwright J and Bundock H (2005) Modified Mercalli intensity maps for the 1906 San Francisco earthquake plotted in shake-map format. US Geological Survey Open-File Report 2005-1135.

Bolt BA (1993) *Abridged Modified Mercalli Intensity Scale, Earthquakes*: W.H. Freeman and Co, Appendix C, 331 pp.

Booker JR (1974) Time dependent strain following faulting of porous medium. *Journal of Geophysical Research* 79: 2037–2044.

Bouchon M, Bouin MP, Karabulut H, Toksoz MN, and Dietrich M (2001) How fast does rupture propagate during an earthquake? New insights from the 1999 Turkey earthquakes. *Geophysical Research Letters* 28: 2723–2726.

Bouchon M and Vallée M (2003) Observation of long supershear rupture during the magnitude 8.1 Kunlunshan earthquake. *Science* 301: 824–826.

Brodsky EE, Mori J, and Fulton PM (2010) Drilling into faults quickly after earthquakes. *Eos, Transactions American Geophysical Union* 91: 237–238.

Brune JN (1968) Seismic moment, seismicity, and rate of slip along major fault zones. *Journal of Geophysical Research* 73: 777–784.

Brune JN, Henyey TL, and Roy RF (1969) Heat flow, stress and rate of slip along San Andreas fault, California. *Journal of Geophysical Research* 74: 3821–3827.

Byerleee JD (1978) Friction of rocks. *Pure and Applied Geophysics* 116: 615–626.

Calais E and Minster JB (1995) GPS detection of ionospheric perturbations following the January 17, 1994, Northridge earthquake. *Geophysical Research Letters* 22: 1045–1048.

Campbell DL (1978) Investigation of the stress-concentration mechanism for intraplate earthquakes. *Geophysical Research Letters* 5: 477–479.

Chapple WM and Forsyth DW (1979) Earthquakes and bending of plates at trenches. *Journal of Geophysical Research* 84: 6729–6749.

Chouet B (1988) Resonance of a fluid-driven crack: Radiation properties and implications for the source of long-period events and harmonic tremor. *Journal of Geophysical Research* 93: 4375–4400.

Chouet B (1996) Long-period volcano seismicity: Its source and use in eruption forecasting. *Nature* 380: 309–316.

Chung W-Y and Kanamori H (1976) Source process and tectonic implications of the Spanish deep-focus earthquake of March 29, 1954. *Physics of the Earth and Planetary Interiors* 13: 85–96.

D'Alessio MA, Blythe AE, and Burgmann R (2003) No frictional heat along the San Gabriel fault, California: Evidence from fission-track thermochronology. *Geology* 31: 541–544. http://dx.doi.org/10.1130/0091-7613.

Delouis B, Nocquet J-M, and Vallée M (2010) Slip distribution of the February 27, 2010 Mw = 8.8 Maule earthquake, central Chile, from static and high-rate GPS, InSAR, and broadband teleseismic data. *Geophysical Research Letters* 37(17), L17305.

Di Toro G, Golbsby DL, and Tullis TE (2004) Friction falls toward zero in quartz rock as slip velocity approaches seismic rates. *Nature* 427: 436–439.

Dieterich JH (1972) Time-dependent friction in rocks. *Journal of Geophysical Research* 77: 3690–3697.

Dieterich JH (1978) Time-dependent friction and the mechanics of stick-slip. *Pure and Applied Geophysics* 116: 790–806.

Dieterich JH (1979a) Modeling of rock friction, 1. Experimental results and constitutive equations. *Journal of Geophysical Research* 84: 2161–2168.

Dieterich JH (1979b) Modeling of rock friction, 2. Simulation of preseismic slip. *Journal of Geophysical Research* 84: 2169–2175.

Dieterich JH (1994) A constitutive law for rate of earthquake production and its application to earthquake clustering. *Journal of Geophysical Research* 99: 2601–2618.

Dragert H, Wang K, and James TS (2001) A silent slip event on the deeper Cascadia subduction interface. *Science* 292: 1525–1528.

Dunham EM and Archuleta RJ (2004) Evidence for a supershear transient during the 2002 Denali fault earthquake. *Bulletin of the Seismological Society of America* 94: S256–S268.

Dunham EM, Belanger D, Cong L, and Kozdon JE (2011) Earthquake ruptures with strongly rate-weakening friction and off-fault plasticity, 2: Nonplanar faults. *Bulletin of the Seismological Society of America* 101(5): 2308–2322. http://dx.doi.org/10.1785/0120100076.

Dziewonski AM, Chou T-A, and Woodhouse JH (1981) Determination of earthquake source parameters from waveform data for studies of global and regional seismicity. *Journal of Geophysical Research* 86: 2825–2852.

Ellsworth WL and Beroza GC (1995) Seismic evidence for an earthquake nucleation phase. *Science* 268: 851–855.

Ellsworth WL, Lindh AG, Prescott WH, and Herd DG (1981) The 1906 San Francisco earthquake and the seismic cycle. In: Simpson DW and Richards PG (eds.)

Earthquake Prediction: An International Review. *Maurice Ewing Series*, vol. 4, pp. 126–140. Washington, DC: AGU.
Engdahl ER, van der Hilst R, and Buland R (1998) Global teleseismic earthquake relocation with improved travel times and procedures for depth determination. *Bulletin of Seismological Society of America* 88: 722–743.
Fedotov SA (1965) Regularities of the distribution of strong earthquakes in Kamchatka, the Kurile islands, and northeast Japan. *Trudy Inst Fiz. Zemli Acad. Nauk. SSSR* 36: 66.
Felzer KR and Brodsky EE (2006) Decay of aftershock activity with distance indicates triggering by dynamic stress. *Nature* 441: 735–738.
Fialko Y and Khazan Y (2005) Fusion by earthquake fault friction: Stick or slip? *Journal of Geophysical Research* 110: B12407. http://dx.doi.org/10.1029/2005JB003869.
Fialko Y, Sandwell D, Agnew D, Simons M, Shearer P, and Minster B (2002) Deformation on nearby faults induced by the 1999 Hector Mine earthquake. *Science* 297: 1858–1862.
Fitch TJ (1972) Plate convergence, transcurrent faults, and internal deformation adjacent to southeast Asia and the western Pacific. *Journal of Geophysical Research* 77: 4432–4460.
Frankel A, Fletcher JB, and Vernon FL (1986) Rupture characteristics and tomographic source imaging of $M_L \sim 3$ earthquakes near Anza, Southern California. *Journal of Geophysical Research* 91(12): 633–650.
Freed AM and Lin J (2001) Delayed triggering of the 1999 Hector Mine earthquake by viscoelastic stress transfer. *Nature* 411: 180–183.
Freund LB (1990) *Dynamic Fracture Mechanics*. Cambridge, UK: Cambridge University Press, 563 pp.
Frohlich C (1989) The nature of deep focus earthquakes. *Annual Review of Earth and Planetary Sciences* 17: 227–254.
Fryer GJ, Watts P, and Pratson LF (2004) Source of the great tsunami of 1 April 1946: A landslide in the upper Aleutian forearc. *Marine Geology* 203: 201–218.
Fukao Y, Hori S, and Ukawa M (1983) A seismological constraint on the depth of the basalt-eclogite transition in a subducting oceanic crust. *Nature* 303: 413–415.
Goldsby DL and Tullis TE (2002) Low frictional strength of quartz rocks at subseismic slip rates. *Geophysical Research Letters* 29: 1844. http://dx.doi.org/10.1029/2002GL015240.
Gordon RG (1998) The plate tectonic approximation: Plate nonrigidity, diffuse plate boundaries, and global plate reconstructions. *Annual Review of Earth and Planetary Sciences* 26: 615–642.
Graves R, Jordan TH, and Callaghan S (2011) Cybershake: A physics-based seismic hazard model for Southern California. *Pure and Applied Geophysics* 168: 367–381.
Guatteri M, Mai PM, and Beroza GC (2004) A pseudo-dynamic approximation to dynamic rupture models for strong ground motion prediction. *Bulletin of the Seismological Society of America* 94(6): 2051–2063. http://dx.doi.org/10.1785/0120040037.
Guatteri M and Spudich P (2000) What can strong-motion data tell us about slip-weakening fault-friction laws. *Bulletin of Seismological Society of America* 90: 98–116. http://dx.doi.org/10.1785/0119990053.
Gutenberg B (1955) Magnitude determination for larger Kern County shocks, 1952; Effects of station azimuths and calculation methods. *California Division of Mines and Geology Bulletin* 171: 171–175.
Gutenberg B and Richter CF (1936) Magnitude and energy of earthquakes. *Science* 83: 183–185.
Gutenberg B and Richter CF (1954) *Seismicity of the Earth and Associated Phenomenon*, 2nd edn. Princeton: Princeton University Press.
Gutenberg B and Richter CF (1942) Earthquake magnitude, intensity, energy, and acceleration. *Bulletin of the Seismological Society of America* 32: 163–191.
Gutenberg B and Richter CF (1956) Magnitude and energy of earthquakes. *Annali di Geofisica* 9: 1–15.
Hagerty MT, Schwartz SY, Garces MA, and Protti M (2000) Analysis of seismic and acoustic observations at Arenal volcano, Costa Rica, 1995–1997. *Journal of Volcanology and Geothermal Research* 101: 27–65.
Hanks TC and Kanamori H (1979) A moment magnitude scale. *Journal of Geophysical Research* 84: 2348–2350.
Harris R (1998) Introduction to special section: Stress triggers, stress shadows, and implications for seismic hazard. *Journal of Geophysical Research* 103: 24347–24358. http://dx.doi.org/10.1029/98JB01576.
Hashimoto C, Noda A, Sagiya T, and Matsu'ura M (2009) Interplate seismogenic zones along the Kuril-Japan trench inferred from GPS data inversion. *Nature Geoscience* 2(2): 141–144.
Healy JH, Rubey WW, Griggs DT, and Raleigh CB (1968) The Denver earthquakes. *Science* 161: 1301–1310.
Heaton TH (1990) Evidence for and implications of self-healing pulses of slip in earthquake rupture. *Physics of the Earth and Planetary Interiors* 64: 1–20.
Hill DP, Reasenberg PA, Michael A, et al. (1993) Seismicity in the western United States remotely triggered by the M 7.4 Landers, California, earthquake of June 28, 1992. *Science* 260: 1617–1623.
Hirose H and Obara K (2005) Repeating short- and long-term slow slip events with deep tremor activity around the Bungo channel region, southwest Japan. *Earth, Planets and Space* 57: 961–972.
Hough SE and Page M (2011) Toward a consistent model for strain accrual and release for the New Madrid Seismic Zone, central United States. *Journal of Geophysical Research* 116. B03311. http://dx.doi.org/10.1029/2010JB007783.
Hubbert MK and Rubey WW (1959) Role of fluid pressure in mechanics of overthrust faulting 1. Mechanics of fluid-filled porous solids and its application to overthrust faulting. *Geological Society of America Bulletin* 70: 115–166.
Huixian L, Housner GW, Lili X, and Duxin H (2002) The great Tangshan earthquake of 1976. Technical Report: CaltechEERL:EERL.2002.001. California Institute of Technology.
Ida Y (1972) Cohesive force across the tip of a longitudinal-shear crack and Griffith's specific surface energy. *Journal of Geophysical Research* 77: 3796–3805.
Ide S and Beroza GC (2001) Does apparent stress vary with earthquake size? *Geophysical Research Letters* 28: 3349–3352.
Ide S, Beroza GC, and McGuire JJ (2005) Imaging earthquake source complexity. In: *Seismic Earth: Array Analysis of Broadband Seismograms*. *Geophysical Monograph Series*, 157. Washington, DC: AGU. http://dx.doi.org/10.1029/157GM08.
Ide S, Beroza GC, Shelly DR, and Uchide T (2007) A scaling law for slow earthquakes. *Nature* 447: 76–79. http://dx.doi.org/10.1038/nature05780.
Ide S, Shelly DR, and Beroza GC (2007) The mechanism of deep low frequency earthquakes: Further evidence that deep non-volcanic tremor is generated by shear slip on the plate interface. *Geophysical Research Letters* 34. L03308. http://dx.doi.org/10.1029/2006GL028890.
Ide S and Takeo M (1997) Determination of constitutive relations of fault slip based on seismic wave analysis. *Journal of Geophysical Research* 102: 27379–27392. http://dx.doi.org/10.1029/97JB02675.
Ishii M, Shearer PM, Houston H, and Vidale J (2005) Extent, duration and speed of the 2004 Sumatra–Andaman earthquake imaged by the Hi-Net array. *Nature* 435: 933–936.
Ito Y and Obara K (2005) Very low frequency earthquakes excited by the 2004 off Kii peninsula earthquakes: A dynamic deformation process in the large accretionary prism. *Earth, Planets and Space* 57: 321–326.
Ito Y and Obara K (2006) Dynamic deformation of the accretionary prism excites very low frequency earthquakes. *Geophysical Research Letters* 33. http://dx.doi.org/10.1029/2005GL025270.
Ito T, Yoshioka S, and Miyazaki Si (2000) Interplate coupling in northeast Japan deduced from inversion analysis of GPS data. *Earth and Planetary Science Letters* 176(1): 117–130. http://dx.doi.org/10.1016/S0012-821X(99)00316-7.
Jeffreys H (1942) On the mechanics of faulting. *Geological Magazine* 79: 291–295.
Johnston AC and Kanter LR (1990) Earthquakes in stable continental crust. *Scientific American* 262(3): 68–75.
Joyner WB and Boore DM (1986) On simulating large earthquakes by Green's function addition of smaller earthquakes. In: *Earthquake Source Mechanics*, pp. 269–274. Washington, DC: American Geophysical Union.
Julian BR (1994) Volcanic tremor: Nonlinear excitation by fluid flow. *Journal of Geophysical Research* 99(11): 859–878.
Julian BR (2002) Seismological detection of slab metamorphism. *Science* 296: 1625–1627.
Julian BR, Miller AD, and Foulger GR (1998) Non-double-couple earthquakes. 1. Theory. *Reviews of Geophysics* 36: 525–549.
Kagan YY and Jackson DD (1991) Seismic gap hypothesis: Ten years after. *Journal of Geophysical Research* 96: 21419–21431.
Kanamori H (1972) Mechanism of tsunami earthquakes. *Physics of the Earth and Planetary Interiors* 6: 346–359.
Kanamori H (1977) The energy release in great earthquakes. *Journal of Geophysical Research* 82: 2981–2987.
Kanamori H and Given JW (1982) Analysis of long-period seismic waves excited by the May 18, 1980, eruption of Mount St. Helens: A terrestrial monopole. *Journal of Geophysical Research* 87: 5422–5432.
Kanamori H and Heaton TH (2000) Microscopic and macroscopic physics of earthquakes. *Geophyiscs and the Complexity of Earthquakes*. *Geophysical Monograph*, 120, pp. 147–163. American Geophysical Union.
Kanamori H and Mori J (1992) Harmonic excitation of mantle Rayleigh waves by the 1991 eruption of Mount Pinatubo, Philippines. *Geophysical Research Letters* 19: 721–724.
Kane MF (1977) Correlation of major eastern earthquake centers with mafic/ultramafic basement masses. US Geological Survey Professional Paper, 1028-O.
Kano Y, Mori J, Fujio R, et al. (2006) Heat signature on the Chelungpu fault associated with the 1999 Chi-Chi, Taiwan earthquake. *Geophysical Research Letters* 33. http://dx.doi.org/10.1029/2006GL026733.

Kelleher JA, Sykes LR, and Oliver J (1973) Possible criteria for predicting earthquake locations and their applications to major plate boundaries of the Pacific and Caribbean. *Journal of Geophysical Research* 78: 2547–2585.

Kirby S, Engdahl ER, and Denlinger R (1996) Intermediate-depth intraslab earthquakes and arc volcanism as physical expressions of crustal and uppermost mantle metamorphism in subducting slabs. In: Bebout GE, et al. (ed.) *Subduction: Top to bottom. Geophysical Monograph,* 96, pp. 195–214. Washington, DC: American Geophysical Union. http://dx.doi.org/10.1029/2001JB000522, 2043.

Kirby SH, Stein S, Okal EA, and Rubie DC (1996) Metastable mantle phase transformations and deep earthquakes in subducting oceanic lithosphere. *Reviews of Geophysics* 34: 261–306.

Kirkpatrick JD, Rowe CD, White JC, and Brodsky EE (2013) Silica gel formation during fault slip: Evidence from the rock record. *Geology* 41: 1015–1018. http://dx.doi.org/10.1130/G34483.1.

Kobayashi N and Nishida K (1998) Continuous excitation of planetary free oscillations by atmospheric disturbances. *Nature* 395: 357–360. http://dx.doi.org/10.1038/26427.

Kostoglodov V, Singh SK, Santiago JA, Larson KM, Lowry AR, and Bilham R (2003) A large silent earthquake in the Guerrero seismic gap, Mexico. *Geophysical Research Letters* 30. http://dx.doi.org/10.1029/2003GL017219.

Kostrov BV and Das S (1988) Principles of earthquake source mechanics. In: *Cambridge Monographs on Mechanics and Applied Mathematics*, p. 286. Cambridge: Cambridge University Press.

Kozdon JE and Dunham EM (2013) Rupture to the trench: Dynamic rupture simulations of the 11 March 2011 Tohoku earthquake. *Bulletin of Seismological Society of America* 103: 1275–1289. http://dx.doi.org/10.1785/0120120136.

Kozdon JE, Dunham EM, and Nordstrom J (2013) Simulation of dynamic earthquake ruptures in complex geometries using high-order finite difference methods. *Journal of Scientific Computing* 55(1): 92–124. http://dx.doi.org/10.1007/s10915-012-9624-5.

Lachenbruch A and Sass J (1980) Heat flow and energetics of the San Andreas fault zone. *Journal of Geophysical Research* 85: 6185–6222.

Lapusta N, Rice JR, Ben-Zion Y, and Zheng G (2000) Elastodynamic analysis for long deformation history with spontaneous rupture episodes on faults with rate- and state-dependent friction. *Journal of Geophysical Research* 105: 23765–23789.

Lay T, Kanamori H, and Ammon CJ (2012) Depth-varying properties of subduction zone megathrust faults. *Journal of Geophysical Research* 117. B04311. http://dx.doi.org/10.1029/2011JB009133.

Lay T, et al. (2010) Teleseismic inversion for rupture process of the 27 February 2010 Chile (Mw 8.8) earthquake. *Geophysical Research Letters* 37(13), L13301.

Leary P, Li YG, and Aki K (1987) Observations and modeling of fault-zone fracture anisotropy I. P., SV, SH travel times. *Geophysical Journal of the Royal Astronomical Society* 91: 461–484.

Li YG, Leary P, Aki K, and Malin P (1990) Seismic trapped modes in the Oroville and San Andreas fault zones. *Science* 249: 763–766.

Li YG, Vidale JE, Aki K, and Xu F (2000) Depth-dependent structure of the Landers fault zone from trapped waves generated by aftershocks. *Journal of Geophysical Research* 105: 6237–6254.

Li YG, Vidale JE, Aki K, Xu F, and Burdette T (1998) Evidence of shallow fault zone strengthening after the 1992 M 7.5 Landers, California, earthquake. *Science* 279: 217–219.

Linde AT, Gladwin MT, Johnston MJS, and Gwyther RL (1996) A slow earthquake sequence on the San Andreas fault. *Nature* 383: 65–69.

Lorito S, et al. (2011) Limited overlap between the seismic gap and coseismic slip of the great 2010 Chile earthquake. *Nature Geoscience* 4(3): 173–177.

Loveless JP and Meade BJ (2010) Geodetic imaging of plate motions, slip rates, and partitioning of deformation in Japan. *Journal of Geophysical Research-Solid Earth* 115(B2), B02410.

Ma S and Andrews DJ (2010) Inelastic off-fault response and three-dimensional dynamics of earthquake rupture on a strike-slip fault. *Journal of Geophysical Research* 115. http://dx.doi.org/10.1029/2009JB006382.

Ma S and Beroza GC (2008) Rupture dynamics on a bi-material interface for dipping faults. *Bulletin of Seismological Society of America* 98: 1642–1658. http://dx.doi.org/10.1785/0120070201.

Mai PM and Beroza GC (2002) A spatial random-field model to characterize complexity in earthquake slip. *Journal of Geophysical Research* 107(B11): 2308. http://dx.doi.org/10.1029/2001JB000588.

Mallet R and Mallet JW (1858)The earthquake catalogue of the British Association, with the discussion, curves, and maps, etc. Transaction of the British Association for the advancement of Science, London.

Marone C and Scholz CH (1988) The depth of seismic faulting and the upper transition from stable to unstable slip regimes. *Geophysical Research Letters* 15: 621–624.

Marza VI (2004) On the death toll of the 1999 Izmit (Turkey) major earthquake. *ESC General Assembly Papers*. Potsdam: European Seismological Commission.

Mazzotti S, Le Pichon X, Henry P, and Miyazaki S-I (2000) Full interseismic locking of the Nankai and Japan-west Kurile subduction zones: An analysis of uniform elastic strain accumulation in Japan constrained by permanent GPS. *Journal of Geophysical Research-Solid Earth* 105(B6): 13159–13177.

McCaffrey R (1992) Oblique plate convergence, slip vectors, and forearc deformation. *Journal of Geophysical Research* 97: 8905–8915.

McCann WR, Nishenko SR, Sykes LR, and Krause J (1979) Seismic gaps and plate tectonics: Seismic potential for major plate boundaries. *Pure and Applied Geophysics* 117: 1082–1147.

McGuire JJ, Zhao L, and Jordan TH (2002) Predominance of unilateral rupture for a global catalog of large earthquakes. *Bulletin of Seismological Society of America* 92: 3309–3317.

McKenzie D and Brune JN (1972) Melting on fault planes during large earthquakes. *Geophysical Journal of the Royal Astronomical Society* 29: 65–78.

McNally KC and McEvilly TV (1977) Velocity contrast across the San Andreas Fault in central California: Small scale variations from P-wave nodal plane distortion. *Bulletin of the Seismological Society of America* 67: 1565–1576.

McNutt SR (2005) Volcanic seismology. *Annual Review of Earth and Planetary Sciences* 32: 461–491. http://dx.doi.org/10.1146/annurev.earth.33.092203.122459.

Meng L, Ampuero J-P, Stock J, Duputel Z, Luo Y, and Tsai VC (2012) Earthquake in a maze: Compressional rupture branching during the 2012 Mw 8.6 Sumatra earthquake. *Science* 337(6095): 724–726. http://dx.doi.org/10.1126/science.1224030.

Michael AJ (1990) Energy constraints on kinematic models of oblique faulting: Loma Prieta versus Parkfield-Coalinga. *Geophysical Research Letters* 17: 1453–1456.

Michael AJ (2011) Random variability explains apparent global clustering of large earthquakes. *Geophysical Research Letters* 38. L21301. http://dx.doi.org/10.1029/2011GL049443.

Miller SA, Collettini C, Chiaraluce L, Cocco M, Barchi M, and Kaus BJ (2004) Aftershocks driven by a high-pressure CO_2 source at depth. *Nature* 427: 724–727.

Miller MM, Melbourne T, Johnson DJ, and Sumner WQ (2002) Periodic slow earthquakes from the Cascadia subduction zone. *Science* 295: 2423.

Mitsui Y, Kato N, Fukuhata Y, and Hirahara K (2012) Megaquake cycle at the Tohoku subduction zone with thermal fluid pressurization near the surface. *Earth and Planetary Science Letters* 325: 21–26. http://dx.doi.org/10.1016/j.epsl.2012.01.026.

Mogi K (1968) Some features of recent seismic activity in and near Japan, 1. *Bulletin of the Earthquake Research Institute Tokyo University* 46: 1225–1236.

Molnar P, Freedman D, and Shih JSF (1979) Lengths of intermediate and deep seismic zones and temperatures in downgoing slabs of lithosphere. *Geophysical Journal International* 56: 41–54.

Montelli R, Nolet G, Dahlen FA, Masters G, Engdahl ER, and Hung S-H (2004) Finite-frequency tomography reveals a variety of plumes in the mantle. *Science* 303: 338–343. http://dx.doi.org/10.1126/science.1092485.

Nadeau RM and Johnson LR (1998) Seismological studies at Parkfield VI: Moment release rates and estimates of source parameters for small repeating earthquakes. *Bulletin of Seismological Society of America* 88: 790–814.

Nason RD and Tocher D (1970) Measurement of movement on the San Andreas fault. In: Mansinha L, Smylie DE, and Beck AE (eds.) *Earthquake Displacement Fields and the Rotation of the Earth*, pp. 246–254. Dordrecht, Netherlands: Reidel.

National Research Council of the National Academies (2003) Living on an Active Earth, perspectives on earthquake science. Committee on the Science of Earthquakes Board on Earth Sciences and Resources Division on Earth and Life Studies. Washington, D.C.: The National Academies Press

Nishimura T, et al. (2004) Temporal change of interplate coupling in northeastern Japan during 1995–2002 estimated from continuous GPS observations. *Geophysical Journal International* 157(2): 901–916.

Nur A and Booker JR (1972) Aftershocks caused by pore fluid flow? *Science* 175: 885–887.

Obara K (2002) Nonvolcanic deep tremor associated with subduction in southwest Japan. *Science* 296: 1679–1681.

Obara K, Hirose H, Yamamizu F, and Kasahara K (2004) *Geophysical Research Letters* 31. L23602. http://dx.doi.org/10.1029/2004GL020848.

Ogata Y (1988) Statistical-models for earthquake occurrences and residual analysis for point-processes. *Journal of the American Statistical Association* 83: 9–27.

Okal EA (1992) Use of the mantle magnitude M_m for the reassessment of the moment of historical earthquakes I: Shallow events. *Pure and Applied Geophysics* 139: 17–57.

Okal E and Stewart LM (1982) Slow earthquakes along oceanic fracture zones: Evidence for asthenospheric flow away from hotspots? *Earth and Planetary Science Letters* 57: 75–87.

Ozawa S, Murakami M, Kaidzu M, and Hatanaka Y (2005) Transient crustal deformation in Tokai region, central Japan, until May 2004. *Earth, Planets and Space* 57: 909–915.
Parsons T, Toda S, Stein RS, Barka A, and Dieterich JH (2000) Heightened odds of large earthquakes near Istanbul: An interaction-based probability calculation. *Science* 288: 661–665.
Perrin G, Rice JR, and Zheng G (1995) Self-healing slip pulse on a frictional surface. *Journal of the Mechanics and Physics of Solids* 43: 1461–1495.
Peyrat S, Olsen KB, and Madariaga R (2001) Dynamic modeling of the 1992 Landers earthquake. *Journal of Geophysical Research* 106(26): 467–482.
Poupinet G and Ellsworth WL (1984) Monitoring velocity variations in the crust using earthquake doublets: An application to the Calaveras fault, California. *Journal of Geophysical Research* 89: 5719–5732.
Quin H (1990) Dynamic stress drop and rupture dynamics of the October 15, 1979 Imperial Valley, California earthquake. *Tectonophysics* 175: 93–117.
Raleigh CB, Healy JH, and Bredehoeft JD (1976) An experiment in earthquake control at Rangely, Colorado. *Science* 191: 1230–1237.
Reinecker J, Heidbach O, Tingay M, Sperner B, and Müller B (2005) The release 2005 of the World Stress Map (available online at www.world-stress-map.org).
Rempel AW and Rice JR (2006) Thermal pressurization and onset of melting in fault zones. *Journal of Geophysical Research* 111. http://dx.doi.org/10.1029/2006JB004314.
Rhie J and Romanowicz B (2004) Excitation of Earth's continuous free oscillations by atmosphere-ocean-seafloor coupling. *Nature* 431: 552–556. http://dx.doi.org/10.1038/nature02942.
Rice JR (1980) The mechanics of earthquake rupture. In: Dziewonski AM and Boschi E (eds.) *Physics of the Earth's Interior*, pp. 555–649. Amsterdam: North-Holland.
Rice JR (2006) Heating and weakening of faults during earthquake slip. *Journal of Geophysical Research* 111. http://dx.doi.org/10.1029/2005JB004006.
Rice JR and Cleary MP (1976) Some basic stress diffusion solutions for fluid-saturated elastic porous media with compressible constituents. *Reviews of Geophysics and Space Physics* 14: 227–241.
Richter CF (1935) An instrumental earthquake magnitude scale. *Bulletin of Seismological Society of America* 25: 1–32.
Richter CF (1958) *Elementary Seismology*. San Francisco, CA: W. H. Freeman and Co, 758 pp.
Risk Management Solutions Inc (1995) *What if the 1923 Earthquake Strikes Again? A Five Prefecture Tokyo Region Scenario*. Menlo Park, CA: Risk Management Solutions Inc., 97 pp.
Rogers G and Dragert H (2003) Episodic tremor and slip on the Cascadia subduction zone: The chatter of silent slip. *Science* 300: 1942–1943.
Rosakis AJ, Samudrala O, and Coker D (1999) Cracks faster than the shear wave speed. *Science* 284: 1337–1340.
Royer AA and Bostock MG (2013) A comparative study of low frequency earthquake templates in northern Cascadia. *Earth and Planetary Science Letters* 42: 247–256. http://dx.doi.org/10.1016/j.epsl.2013.08.040.
Royer J-Y and Gordon RG (1997) The motion between the Capricorn and Australian Plates. *Science* 277: 1268–1274.
Rubin AM, Gillard D, and Got J-L (1999) Streaks of microearthquakes along creeping faults. *Nature* 400: 635–641.
Ruegg JC, et al. (2009) Interseismic strain accumulation measured by GPS in the seismic gap between constitución and concepción in Chile. *Physics of the Earth and Planetary Interiors* 175(1–2): 78–85.
Ruina A (1983) Slip instability and state variable friction laws. *Journal of Geophysical Research* 88(10): 359–370.
Savage JC (1983) A dislocation model of strain accumulation and release at a subduction zone. *Journal of Geophysical Research* 88: 4984–4996.
Sbar ML and Sykes LR (1973) Contemporary compressive stress and seismicity in eastern North America: An example of intra-plate tectonics. *Geological Society of America Bulletin* 84: 1861–1882.
Schaff DP, Bokelmann GHR, Beroza GC, Waldhauser F, and Ellsworth WL (2002) High resolution image of Calaveras Fault seismicity. *Journal of Geophysical Research* 107(B9): 2186. http://dx.doi.org/10.1029/2001JB000633.
Schaff DP, Bokelmann GHR, Ellsworth WL, Zanzerkia E, Waldhauser F, and Beroza GC (2004) Optimizing correlation techniques for improved earthquake location. *Bulletin of Seismological Society of America* 94: 705–721.
Scholz CH (1968) Microfractures, aftershocks, and seismicity. *Bulletin of Seismological Society of America* 58: 1117–1130.
Scholz CH (1989) Mechanics of faulting. *Annual Review of Earth and Planetary Sciences* 17: 309–334.
Scholz CH (1990) *The Mechanics of Earthquakes and Faulting*. New York: Cambridge University Press.
Scholz CH (2000) Evidence for a strong San Andreas Fault. *Geology* 28: 163–166. http://dx.doi.org/10.1130/0091-7613.
Scholz CH, Molnar P, and Johnson T (1972) Frictional sliding of granite and earthquake mechanism implications. *Journal of Geophysical Research* 77: 6392–6406.
Seno T and Yamasaki T (2003) Low-frequency tremors, intraslab and interplate earthquakes in Southwest Japan—From a viewpoint of slab dehydration. *Geophysical Research Letters* 30. http://dx.doi.org/10.1029/2003GL018349.
Shearer PM, Prieto GA, and Hauksson E (2006) Comprehensive analysis of earthquake source spectra in southern California. *Journal of Geophysical Research* 111: B06303. http://dx.doi.org/10.1029/2005JB003979.
Shearer PM and Stark PB (2012) The global risk of big earthquakes has not recently increased. *Proceedings of the National Academy of Sciences* 109: 717–721. http://dx.doi.org/10.1073/pnas.1118525109.
Shelly DR, Beroza GC, and Ide S (2007) Non-volcanic tremor and low frequency earthquake swarms. *Nature* 446. http://dx.doi.org/10.1038/nature05666.
Shelly DR, Beroza GC, Ide S, and Nakamula S (2006) Low-frequency earthquakes in Shikoku, Japan, and their relationship to episodic tremor and slip. *Nature* 442: 188–191.
Shibazaki B, Matsuzawa T, Tsutsumi A, Ujiie K, Hasegawa A, and Ito Y (2011) 3D modeling of the cycle of a great Tohoku-Oki earthquake, considering frictional behavior at low to high slip velocities. *Geophysical Research Letters* 38: L21305. http://dx.doi.org/10.1029/2011GL049308.
Sibson RH (1973) Interaction between temperature and pore-fluid pressure during earthquake faulting—A mechanism for partial or total stress relief. *Nature* 243: 66–68.
Sibson RH (1981) In: Simpson DW and Richards PG (eds.) *Earthquake Prediction: An International Review. Maurice Ewing Series*, 4: pp. 593–603. Washington, DC: AGU Press.
Singh SK, Mena E, and Castro R (1988) Some aspects of source characteristics of the 19 September 1985 Michoacan earthquake and ground motion amplification in and near Mexico City from strong motion data. *Bulletin of Seismological Society of America* 78: 451–477.
Singh SK and Ordaz M (1994) Seismic energy release in Mexican subduction zone earthquakes. *Bulletin of Seismological Society of America* 84: 1533–1550.
Somerville PG, Smith NF, Graves RW, and Abrahamson NA (1997) Modification of empirical strong ground motion attenuation relations to include the amplitude and duration effects of rupture directivity. *Seismological Research Letters* 68: 199–222.
Stein RS (1999) The role of stress transfer in earthquake occurrence. *Nature* 402: 605–609. http://dx.doi.org/10.1038/45144.
Stein RS, Barka AA, and Dieterich JH (1997) Progressive failure on the North Anatolian fault since 1939 by earthquake stress triggering. *Geophysical Journal International* 128: 594–604.
Stirling MW and Petersen M (2006) Comparison of the historical record of earthquake hazard with seismic-hazard models for New Zealand and the continental United States. *Bulletin of Seismological Society of America* 96: 1978–1994.
Stirling M, Wesnousky SG, and Shimizaki K (1996) Fault trace complexity, cumulative slip, and the shape of the magnitude-frequency distribution for strike-slip faults—A global survey. *Geophysical Journal International* 124: 833–868.
Suda N, Nawa K, and Fukao Y (1998) Earth's background free oscillations. *Science* 279: 2089–2091. http://dx.doi.org/10.1126/science.279.5359.2089.
Suwa Y, Miura S, Hasegawa A, Sato T, and Tachibana K (2006) Interplate coupling beneath NE Japan inferred from three-dimensional displacement field. *Journal of Geophysical Research-Solid Earth* 111(B4), B04402.
Sykes LR (1967) Mechanism of earthquakes and nature of faulting on the mid-oceanic ridges. *Journal of Geophysical Research* 72: 2131–2153.
Tanaka H, Chen W-M, and Wang C-Y (2006) Frictional heat from faulting of the 1999 Chi-Chi, Taiwan, earthquake. *Geophysical Research Letters* 33: L16316. http://dx.doi.org/10.1029/2006GL026673.
Tauzin B, Debayle E, Quantin C, and Coltice N (2013) Seismoacoustic coupling induced by the breakup of the 15 February 2013 Chelyabinsk meteor. *Geophysical Research Letters* 40: 3522–3526. http://dx.doi.org/10.1002/grl.506832013.
Titus SJ, DeMets C, and Tikoff B (2006) Thirty-five-year creep rates for the creeping section of the San Andreas Fault and the effects of the 2004 Parkfield earthquake; constraints from alignment arrays, continuous global positioning system, and creepmeters. *Bulletin of Seismological Society of America* 96: S250–S268. http://dx.doi.org/10.1785/0120050811.
Toda S, Stein RS, Beroza GC, and Marsan D (2012) Aftershocks halted by static stress shadows. *Nature Geoscience* 5: 410–413. http://dx.doi.org/10.1038/ngeo1465.
Toksöz MN, Shakal AF, and Michael AJ (1979) Space-time migration of earthquakes along the North Anatolian fault zone and seismic gaps. *Pure and Applied Geophysics* 117: 1258–1270.
Toppozada TR (1975) Earthquake magnitude as a function of intensity data in California and western Nevada. *Bulletin of Seismological Society of America* 65: 1223–1238.

Toppozada TR, Real CR, and Park DL (1981) Preparation of isoseismal maps and summaries of reported effects for pre-1900 California earthquakes. California. Division of Mines and Geology Open File Pert. 81-11 SAC, 182 pp.

Tse ST and Rice JR (1986) Crustal earthquake instability in relation to the depth variation of frictional slip properties. *Journal of Geophysical Research* 91: 9452–9472.

Unsworth MJ, Malin PE, Egbert GD, and Booker JR (1997) Internal Structure of the San Andreas fault at Parkfield, California. *Geology* 25: 359–362.

Utsu T (1961) A statistical study on the occurrence of aftershocks. *Geophysical Magazine* 30: 521–605.

Vanek J, Zapotek A, Karnik V, et al. (1962) Standardization of magnitude scales. *Izvestiya Akademii Nauk SSSR Seriya Geofizicheskaya* 2: 153–158.

Vigny C, et al. (2011) The 2010 Mw 8.8 Maule megathrust earthquake of central Chile, monitored by GPS. *Science* 332(6036): 1417–1421.

Wald DJ and Heaton TH (1994) Spatial and temporal distribution of slip for the 1992 Landers, California earthquake. *Bulletin of the Seismological Society of America* 84: 668–691.

Wald DJ, Quitoriano V, Heaton T, and Kanamori H (1999) Relationships between peak ground acceleration, peak ground velocity, and modified Mercalli intensity in California. *Earthquake Spectra* 15: 557–564.

Waldhauser F and Ellsworth WL (2000) A double-difference earthquake location algorithm: Method and application to the northern Hayward Fault, California. *Bulletin of Seismological Society of America* 90: 1353–1368.

Waldhauser F, Ellsworth WL, Schaff DP, and Cole A (2004) Streaks, multiplets, and holes: High-resolution spatio-temporal behavior of Parkfield seismicity. *Geophysical Research Letters* 31, L18608.

Walter WR, Rodgers AJ, Mayeda K, Myers SC, Pasyanos M, and Denny M (1998) Preliminary regional seismic analysis of nuclear explosions and earthquakes in Southwest Asia. In: Proceedings of the 20th Annual Seismic Research Symposium on Monitoring a Comprehensive Test Ban Treaty, pp. 442–452. Department of Defense and Department of Energy Report.

Ward SN (1995) Area-based tests of long-term seismic hazard predictions. *Bulletin of Seismological Society of America* 85: 1285–1298.

Wech AG and Creager KC (2007) Cascadia tremor polarization evidence for plate interface slip. *Geophysical Research Letters* 34: L22306. http://dx.doi.org/10.1029/2007GL031167.

Wesnousky SG (1994) The Gutenberg–Richter or characteristic earthquake distribution, which is it? *Bulletin of Seismological Society of America* 84: 1940–1959.

Wesson RL (1988) Dynamics of fault creep. *Journal of Geophysical Research* 93: 8929–8951.

West M, Sanchez JJ, and McNutt SR (2005) Periodically triggered seismicity at Mount Wrangell, Alaska, after the Sumatra earthquake. *Science* 308: 1144–1146. http://dx.doi.org/10.1126/science.1112462.

Wibberley CAJ and Shimamoto T (2003) Internal structure and permeability of major strike-slip fault zones: The median tectonic line in mie prefecture, Southwest Japan. *Journal of Structural Geology* 25: 59–78.

Widmer R and Zürn W (1992) Bichromatic excitation of long-period Rayleigh and air waves by the Mount Pinatubo and El Chichon volcanic eruptions. *Geophysical Research Letters* 19: 765–768.

Wong T-F (1982) Shear fracture energy of Westerly granite from postfailure behavior. *Journal of Geophysical Research* 87: 990–1000.

Working Group on California Earthquake Probabilities (2002) Earthquake probabilities in the San Francisco Bay region: 2002–2031. USGS Open File Report 03-214.

Wu J, Wallace T, and Beck S (1995) A very broadband study of the 1994 deep Bolivia earthquake sequence. *Geophysical Research Letters* 22: 2237–2240.

Xue L, Li H-B, and Brodsky EE (2013) Continuous permeability measurements record healing inside the Wenchuan earthquake fault zone. *Science* 340: 1555–1559. http://dx.doi.org/10.1126/science.1237237.

Yong C, Tsoi K-L, Feibi C, Henhuan GZ, Qijia Z, and Zhangli C (1988) *The Great Tangshan Earthquake of 1976: An Anatomy of Disaster*. Oxford: Pergamon Press, 153 pp.

Yu G, Wesnousky SG, and Ekstrom G (1993) Slip partitioning along major convergent plate boundaries. *Pure and Applied Geophysics* 140: 183–210.

Zhang M and Wen L (2013) High-precision location and yield of North Korea's 2013 nuclear test. *Geophysical Research Letters* 40: 2941–2946. http://dx.doi.org/10.1002/.50607.

Zhao L-F, Xie X-B, Wang W-M, and Yao Z-X (2012) Yield estimation of the 25 May, 2009 North Korean nuclear explosion. *Bulletin of Seismological Society of America* 102: 467–478. http://dx.doi.org/10.1785/0120110163.

Zoback ML, Zoback MD, Adams J, et al. (1989) Global patterns of tectonic stress. *Nature* 341: 291–298. http://dx.doi.org/10.1038/341291a0.

Zoback MD, Zoback ML, and Mount V (1987) New evidence on the state of stress on the San Andreas fault system. *Science* 238: 1105–1111.

4.02 Seismic Source Theory

R Madariaga, Laboratoire de Géologie Ecole Normale Supérieure, Paris, France

© 2015 Elsevier B.V. All rights reserved.

4.02.1	**Introduction**	51
4.02.2	**Seismic Wave Radiation from a Point Force: The Green Function**	52
4.02.2.1	Seismic Radiation from a Point Source	52
4.02.2.2	Far-Field Body Waves Radiated by a Point Force	52
4.02.2.3	The Near Field of a Point Force	53
4.02.2.4	Energy Flow from Point Force Sources	53
4.02.2.5	The Green Tensor for a Point Force	54
4.02.3	**Moment Tensor Sources**	54
4.02.3.1	Radiation from a Point Moment Tensor Source	55
4.02.3.2	A More General View of Moment Tensors	55
4.02.3.3	Moment Tensor Equivalent of a Fault	56
4.02.3.4	Eigenvalues and Eigenvectors of the Moment Tensor	57
4.02.3.5	Seismic Radiation from Moment Tensor Sources in the Spectral Domain	57
4.02.3.6	Seismic Energy Radiated by Point Moment Tensor Sources	58
4.02.3.7	More Realistic Radiation Models	58
4.02.4	**Finite Source Models**	59
4.02.4.1	The Kinematic Dislocation Model	59
4.02.4.1.1	Haskell's rectangular fault model	60
4.02.4.2	The Circular Fault Model	61
4.02.4.2.1	Kostrov's self-similar circular crack	61
4.02.4.2.2	The kinematic circular source model of Sato and Hirasawa	62
4.02.4.3	Generalization of Kinematic Models and the Isochrone Method	62
4.02.5	**Crack Models of Seismic Sources**	63
4.02.5.1	Rupture Front Mechanics	64
4.02.5.2	Stress and Velocity Intensity	65
4.02.5.3	Energy Flow into the Rupture Front	65
4.02.5.4	The Circular Crack	65
4.02.5.5	The Dynamic Circular Fault in a Homogenous Medium	68
4.02.6	**Conclusions**	69
Acknowledgments		70
References		70

4.02.1 Introduction

Earthquake source dynamics provides key elements for the prediction of ground motion and to understand the physics of earthquake initiation, propagation, and healing. The simplest possible model of seismic source is that of a point source buried in an elastic half-space. The development of a proper model of the seismic source took more than 50 years since the first efforts by Nakano (1923) and colleagues in Japan. Earthquakes were initially modeled as simple explosions, then as the result of the displacement of conical surfaces, and finally as the result of fast transformational strains inside a sphere. In the early 1950s, it was recognized that P-waves radiated by earthquakes presented a spatial distribution similar to that produced by single couples of forces, but it was very soon recognized that this type of source could not explain S-wave radiation (Honda, 1962). The next level of complexity was to introduce a double-couple source, a source without resultant force nor moment. The physical origin of the double-couple model was established in the early 1960s thanks to the observational work of numerous seismologists and the crucial theoretical breakthrough of Maruyama (1963) and Burridge and Knopoff (1964) who proved that a fault in an elastic model was equivalent to a double-couple source.

In this chapter, we shall review what we believe are the essential results obtained in the field of kinematic earthquake rupture to date. In **Section 4.02.2**, we review the classical point source model of elastic wave radiation and establish some basic general properties of energy radiation by that source. In **Section 4.02.3**, we discuss the now classical seismic moment tensor source. In **Section 4.02.4**, we discuss extended kinematic sources including the simple rectangular fault model proposed by Haskell (1964, 1966) and a circular model that tries to capture some essential features of crack models. **Section 4.02.5** introduces crack models without friction as models of shear faulting in the earth. This will help establish some basic results that are useful in the study of dynamic models of the earthquake source.

 http://dx.doi.org/10.1016/B978-0-444-53802-4.00070-1

4.02.2 Seismic Wave Radiation from a Point Force: The Green Function

There are many ways of solving the elastic wave equation for different types of initial conditions, boundary conditions, sources, etc. Each of these methods requires a specific approach for every different problem that we would need to study. Ideally, we would like, however, to find a general solution method that would allow us to solve any problem by a simple method. The basic building block of such a general solution method is the Green function, the solution of the following elementary problem: find the radiation from a point source in a finite heterogeneous elastic medium. For simplicity, we consider first the particular case of a homogeneous elastic isotropic medium, for which we know how to calculate the Green function. This will let us establish a general framework for studying more elaborate source models.

4.02.2.1 Seismic Radiation from a Point Source

The simplest possible source of elastic waves is a point force of arbitrary orientation located inside an infinite homogeneous, isotropic elastic body of density ρ, and elastic constants λ and μ. Let $\alpha=\sqrt{(\lambda+2\mu)/\rho}$ and $\beta=\sqrt{\mu/\rho}$ be the P- and S-wave speeds, respectively. Let us note $\mathbf{u}(\mathbf{x},t)$, the particle displacement vector. We have to find the solution of the elastodynamic wave equation:

$$\rho\frac{\partial^2}{\partial t^2}\mathbf{u}(\mathbf{x},t)=(\lambda+\mu)\nabla(\nabla\cdot\mathbf{u}(\mathbf{x},t))+\mu\nabla^2\mathbf{u}(\mathbf{x},t)+\mathbf{f}(\mathbf{x},t) \quad [1]$$

under homogeneous initial conditions, that is, $\mathbf{u}(\mathbf{x},0)=\dot{\mathbf{u}}(\mathbf{x},0)=0$, and the appropriate radiation conditions at infinity. In [1], $\mathbf{f}$ is a general distribution of force density as a function of position and time. For a point force of arbitrary orientation located at a point $\mathbf{x}_0$, the body force distribution is

$$\mathbf{f}(\mathbf{x},t)=\mathbf{f}s(t)\delta(\mathbf{x}-\mathbf{x}_0) \quad [2]$$

where $s(t)$ is the source time function, the variation of the amplitude of the force as a function of time. And $\mathbf{f}$ is a unit vector in the direction of the unit force.

The solution of eqn [1] is easier to obtain in the Fourier transformed domain. As is usual in seismology, we use the following definition of the Fourier transform and its inverse:

$$\begin{aligned}\widetilde{u}(x,\omega)&=\int_{-\infty}^{\infty}u(x,t)\mathrm{e}^{-i\omega t}\mathrm{d}t\\ u(x,t)&=\frac{1}{2\pi}\int_{-\infty}^{\infty}\widetilde{u}(x,\omega)\mathrm{e}^{-i\omega t}\mathrm{d}\omega\end{aligned} \quad [3]$$

Here and in the following, we will note Fourier transform with a tilde.

After some lengthy work see, for example, Achenbach (1975), we find the Green function in the Fourier domain:

$$\begin{aligned}\widetilde{\mathbf{u}}(R,\omega)=&\frac{1}{4\pi\rho}\left[\mathbf{f}\cdot\nabla\nabla\left(\frac{1}{R}\right)\right]\frac{\widetilde{s}(\omega)}{\omega^2}\left[-\left(1+\frac{i\omega R}{\alpha}\right)\mathrm{e}^{-i\omega R/\alpha}\right.\\&\left.+\left(1+\frac{i\omega R}{\beta}\right)\mathrm{e}^{-i\omega R/\beta}\right]\\&+\frac{1}{4\pi\rho\alpha^2}\frac{1}{R}(\mathbf{f}\cdot\nabla R)\nabla R\widetilde{s}(\omega)\mathrm{e}^{-i\omega R/\alpha}\\&+\frac{1}{4\pi\rho\beta^2}\frac{1}{R}[\mathbf{f}-(\mathbf{f}\cdot\nabla R)\nabla R]\widetilde{s}(\omega)\mathrm{e}^{-i\omega R/\beta}\end{aligned} \quad [4]$$

where $R=\|\mathbf{x}-\mathbf{x}_0\|$ is the distance from the source to the observation point. Using the inverse Fourier transform, we can transform eqn [4] to the time domain to obtain the final result

$$\begin{aligned}\mathbf{u}(R,t)=&\frac{1}{4\pi\rho}[\mathbf{f}\cdot\nabla\nabla(1R)]\int_{R/\alpha}^{\min(t,R/\beta)}\tau s(t-\tau)\mathrm{d}\tau\\&+\frac{1}{4\pi\rho\alpha^2}\frac{1}{R}[(\mathbf{f}\cdot\nabla R)\nabla R]s(t-R/\alpha)\\&+\frac{1}{4\pi\rho\beta^2}\frac{1}{R}[\mathbf{f}-(\mathbf{f}\cdot\nabla R)\nabla R]s(t-R/\beta)\end{aligned} \quad [5]$$

This complicated-looking expression can be better understood considering each of its terms separately. The first line is the near field, which comprises all the terms that decrease at long distance from the source faster than R^{-1}. The last two lines are the far-field P and S spherical waves that decrease with distance like R^{-1}.

4.02.2.2 Far-Field Body Waves Radiated by a Point Force

Much of the practical work of seismology is done in the far field, at distances of several wavelengths from the source. When the distance R is large, only the last two terms in [5] are important. Under what conditions can we neglect the first term of that expression with respect to the last two? For that purpose, we notice that in [4], R appears always in the particular combination $\omega R/\alpha$ or $\omega R/\beta$. Clearly, these two ratios determine the far-field conditions. Since $\alpha>\beta$, we conclude that the far field is defined by

$$\frac{\omega R}{\alpha}\gg 1 \quad \text{or} \quad \frac{R}{\lambda}\ll 1$$

where $\lambda=2\pi\alpha/\omega$ is the wavelength of a P-wave of circular frequency ω. The condition for the far field depends therefore on the characteristic frequency or wavelength of the radiation. Thus, depending on the frequency content of the signal $\widetilde{S}(\omega)$, we will be in the far field for high-frequency waves, but we may be in the near field for the low-frequency components.

The far-field radiation from a point force is usually written in the following, shorter form:

$$\begin{aligned}\mathbf{u}^{\mathrm{P}}_{\mathrm{FF}}(R,t)&=\frac{1}{4\pi\rho\alpha^2}\frac{1}{R}\mathfrak{R}^{\mathrm{P}}s(t-R/\alpha)\\ \mathbf{u}^{\mathrm{S}}_{\mathrm{FF}}(R,t)&=\frac{1}{4\pi\rho\beta^2}\frac{1}{R}\mathfrak{R}^{\mathrm{S}}s(t-R/\beta)\end{aligned} \quad [6]$$

where $\mathfrak{R}^{\mathrm{P}}$ and $\mathfrak{R}^{\mathrm{S}}$ are the radiation patterns of P- and S-waves, respectively. Noting that $\nabla R=\mathbf{e}_R$, the unit vector in the radial direction, we can write the radiation patterns in the following simplified form: $\mathfrak{R}^{\mathrm{P}}=f_R\mathbf{e}_R$ and $\mathfrak{R}^{\mathrm{S}}=\mathbf{f}_{\mathrm{T}}=\mathbf{f}-f_{\mathrm{R}}\mathbf{e}_{\mathrm{R}}$ where f_{R} is the radial component of the point force $\mathbf{f}$ and $\mathbf{f}_{\mathrm{T}}$ its transverse component.

Thus, in the far field of a point force, P-waves propagate the radial component of the point force, while the S-waves propagate information about the transverse component of the S-wave. Expressing the amplitude of the radial and transverse component of $\mathbf{f}$ in terms of the azimuth θ of the ray with respect to the applied force, we can rewrite the radiation patterns in the simpler form

$$\Re^{\mathrm{P}} = \cos\theta \mathbf{e}_{\mathrm{R}}, \quad \Re^{\mathrm{S}} = \sin\theta \mathbf{e}_{\Gamma} \qquad [7]$$

As we could expect from the natural symmetry of the problem, the radiation patterns are axially symmetrical about the axis of the point force. P-waves from a point force have a typical dipolar radiation pattern, while S-waves have a toroidal (doughnut-shaped) distribution of amplitudes.

4.02.2.3 The Near Field of a Point Force

When $\omega R/\alpha$ is not large compared to one, all the terms in eqns [4] and [5] are of equal importance. In fact, both far- and near-field terms are of the same order of magnitude near the point source. In order to find the small R behavior, it is preferable to go back to the frequency domain expression [4]. When $R \to 0$, the term in brackets in the first line tends to zero. In order to calculate the near-field behavior, we have to expand the exponentials to order R^2. After some algebra, we find

$$\mathbf{u}(R,t) = \frac{1}{8\pi\rho\beta^2}\frac{1}{R}\left[(\mathbf{f}\cdot\nabla R)\nabla R\left(1-\beta^2/\alpha^2\right) + \mathbf{f}\left(1+\beta^2/\alpha^2\right)\right]s(t) \qquad [8]$$

This is the product of the source time function $s(t)$ with the static displacement produced by a point force of orientation $\mathbf{f}$. This is one of the most important results of static elasticity and is frequently referred to as the Kelvin or Somigliana solution (Aki and Richards, 2002).

The result [8] is quite interesting and somewhat unexpected. The radiation from a point source decays like R^{-1} in the near field, exactly like the far-field terms. This result has been remarked and extensively used in the formulation of regularized boundary integral equations for elastodynamics (Fukuyama and Madariaga, 1995; Hirose and Achenbach, 1989).

4.02.2.4 Energy Flow from Point Force Sources

A very important issue in seismology is the amount of energy radiated by seismic sources. The flow of energy across any surface that encloses the point source must be the same, so that seismic energy is defined for any arbitrary surface. Let us take the scalar product of eqn [1] with the particle velocity $\dot{\mathbf{u}}$ and integrate on a volume V enclosing the single point source located at $\mathbf{x}_0$:

$$\int_V \rho\dot{\mathbf{u}}_i\ddot{\mathbf{u}}_i \mathrm{d}V = \int_V \sigma_{ij,j}\dot{u}_i \mathrm{d}V \int_V f_i\dot{u}_i \mathrm{d}V \qquad [9]$$

where we use dots to indicate time derivatives and the summation convention on repeated indices. In [9], we have rewritten the right-hand side of [1] in terms of the stresses $\sigma_{ij} = \lambda\varepsilon_{ii}\delta_{ij} + 2\mu\varepsilon_{ij}$, where the strains are $\varepsilon_{ij} = 1/2(\partial_j u_i + \partial_j u_i)$.

Using $\sigma_{ij,j}\dot{u}_i = \left(\sigma_{ij}\dot{u}_i\right)_{,j} - \sigma_{ij}\dot{\varepsilon}_{ij}$ and Gauss' theorem, we get the energy flow identity

$$\frac{\mathrm{d}}{\mathrm{d}t}(K(t)+U(t)) = \int_S \sigma_{ij}\dot{u}_i n_j \mathrm{d}S + \int_U f_i\dot{u}_i \mathrm{d}V \qquad [10]$$

where $\mathbf{n}$ is the outward normal to the surface S (see Figure 1). In [10], K is the kinetic energy contained in volume V at time t:

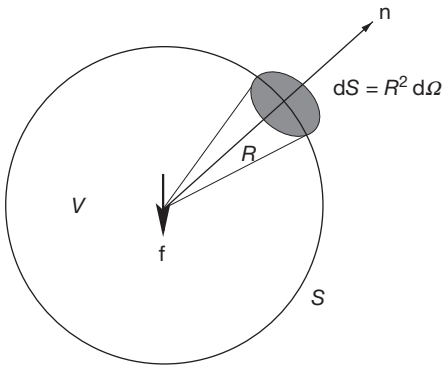


Figure 1 Geometry for computing radiated energy from a point source. The source is at the origin and the observer at a position defined by the spherical coordinates R, θ, ϕ, distance, polar angle, and azimuth.

$$K(t) = \frac{1}{2}\int_V \rho\dot{\mathbf{u}}^2 \mathrm{d}V \qquad [11]$$

while

$$U(t) = \frac{1}{2}\int_V \left[\lambda(\nabla\cdot\mathbf{u})^2 + 2\mu\varepsilon_{ij}\varepsilon_{ij}\right]\mathrm{d}V \qquad [12]$$

is the strain energy change inside the same volume. The last term in [10] is the rate of work of the force against elastic displacement. Equation [10] is the basic energy conservation statement for elastic sources. It says that the rate of energy change inside the body V is equal to the rate of work of the sources $\mathbf{f}$ plus the inward energy flow across the boundary S. This energy balance equation can be extended to study energy flow for moment tensor sources and for cracks (Kostrov and Das, 1988).

Let us note that in [10], energy flows into the body through the surface S. In seismology, however, we are interested in the radiated seismic energy, which is the energy that flows out of the elastic body. The radiated seismic energy until a certain time t is

$$\begin{aligned} E_{\mathrm{s}}(t) &= -\int_0^t \mathrm{d}t \int_S \sigma_{ij}\dot{u}_i n_j \mathrm{d}S \\ &= -K(t) - \Delta U(t) + \int_0^t \mathrm{d}t \int_V f_i\dot{u}_i \mathrm{d}V \end{aligned} \qquad [13]$$

where $\Delta U(t)$ is the strain energy change inside the elastic body from time $t=0$ to t. If t is sufficiently long, so that all motion inside the body has ceased, $K(t) \to 0$ and we get the simplest possible expression

$$E_s = E_s(\infty) = -\Delta U + \int_0^\infty \mathrm{d}t \int_V f_i\dot{u}_i \mathrm{d}V \qquad [14]$$

Thus, total energy radiation is equal to the decrease in internal energy plus the work of the sources against the elastic deformation. Both terms in [14] contribute to seismic radiation as we will discuss in more detail for moment tensor sources and cracks.

Although we can use [14] to compute the seismic energy, it is easier to evaluate it directly from the first term in [13] assuming that S is very far from the source, so that far-field approximation [6] can be used in [13]. Consider, as shown in Figure 1, a cone of rays of cross section $\mathrm{d}\Omega$ issued from the source around the direction θ, ϕ. The energy crossing a section of this ray beam at distance R from the source per unit time is given by the energy

flow per unit solid angle $\dot{e}_s = \sigma_{ij}\dot{u}_i n_j R^2$, where σ_{ij} is the stress, $\dot{u}_i$ the particle velocity, and $\mathbf{n}$ the normal to the surface $\mathrm{d}S = R^2\mathrm{d}\Omega$. We now use [6] in order to compute σ_{ij} and $\dot{u}$. By straightforward differentiation and keeping only terms of order $1/R$ with distance, we get $\sigma_{ij}n_j \approx \rho c\dot{u}_i$ where ρc is the wave impedance and c the appropriate wave speed. The energy flow rate per unit solid angle for each type of wave is then

$$\dot{e}_s(t) = \rho c R^2 \dot{u}^2(R,t) \qquad [15]$$

Using the far-field approximation for the velocity derived from [6] and integrating around the source for the complete duration of the source, we get the total energy flow associated with P- and S-waves:

$$\begin{aligned} E_s^P &= \frac{1}{4\pi\rho\alpha^3}\left\langle \Re^P \right\rangle_2 \int_0^\infty s^{-2}(t)\mathrm{d}t \quad \text{for P waves} \\ E_s^S &= \frac{1}{4\pi\rho\beta^3}\left\langle \Re^S \right\rangle_2 \int_0^\infty s^{-2}(t)\mathrm{d}t \quad \text{for S waves} \end{aligned} \qquad [16]$$

where $\langle \Re^c \rangle_2 = 1/(4\pi)\int_\Omega (\Re^c)^2 \mathrm{d}\Omega$ is the mean squared radiation pattern for wave $c = \{\mathrm{P}, \mathrm{S}\}$. Since the radiation patterns are the simple sinusoidal functions listed in [7], the mean square radiation patterns are 1/3 for P-waves and 2/3 for the sum of the two components of S-waves. In [16], we assumed that $\dot{s}(t) = 0$ for $t < 0$. Finally, it is not difficult to verify that, since $\dot{s}$ has units of force rate, E_s and E_p have units of energy. Noting that in the Earth, α is roughly $\sqrt{3\beta}$ so that $\alpha^3 \cong 5\beta^3$; the amount of energy carried by the S-waves emitted by a point force of source time function $s(t)$ is close to ten times that carried by P-waves. For double-couple sources, this ratio is much larger.

4.02.2.5 The Green Tensor for a Point Force

The Green function is a tensor formed by the waves radiated from a set of three point forces aligned in the direction of each coordinate axis. For an arbitrary force of direction $\mathbf{f}$, located at point $\mathbf{x}_0$ and source function $s(t)$, we define the Green tensor for elastic waves by

$$\mathbf{u}(\mathbf{x},t) = G(\mathbf{x},t|\mathbf{x}_0,0)\cdot\mathbf{f} * s(t)$$

where the star indicates time-domain convolution.

We can also write this expression in the usual index notation

$$u_i(\mathbf{x},t) = \sum_j G_{ij}(\mathbf{x},t|\mathbf{x}_0,0) f_j * s(t)$$

in the time domain or

$$\widetilde{u}_i(\mathbf{x},\omega) = \sum_j \widetilde{G}_{ij}(\mathbf{x}|\mathbf{x}_0,\omega) f_j \widetilde{s}(\omega)$$

in the frequency domain.

The Green function itself can be easily obtained from the radiation from a point force [6]

$$\begin{aligned} G_{ij}(\mathbf{x},t|\mathbf{x}_0,0) = {} & \frac{1}{4\pi\rho}\left(\frac{1}{R}\right)_{,ij} t[\mathrm{H}(t-R/\alpha) - H(t-R/\beta)] \\ & + \frac{1}{4\pi\rho\alpha^2}\frac{1}{R}(R_{,i}R_{,j})\delta(t-R/\alpha) \\ & + \frac{1}{4\pi\rho\beta^2}\frac{1}{R}\left[\delta_{ij} - R_{,i}R_{,j}\right]\delta(t-R/\beta) \end{aligned} \qquad [17]$$

Here, $\delta(t)$ is Dirac's delta, δ_{ij} is Kronecker's delta, and the comma indicates derivative with respect to the component that follows it.

Similarly, in the frequency domain,

$$\begin{aligned} \widetilde{G}_{ij}(\mathbf{x}|\mathbf{x}_0,\omega) = {} & \frac{1}{4\pi\rho}\left(\frac{1}{R}\right)_{,ij}\frac{1}{\omega^2}\left[-(1+i\omega R\alpha)\mathrm{e}^{-i\omega R/\alpha}\right. \\ & \left. + (1+i\omega R\beta)\mathrm{e}^{-i\omega R/\beta}\right] \\ & + \frac{1}{4\pi\rho\alpha^2}\frac{1}{R}(R_{,i}R_{,j})\mathrm{e}^{-i\omega R/\alpha} \\ & + \frac{1}{4\pi\rho\beta^2}\frac{1}{R}\left[\delta_{ij} - R_{,i}R_{,j}\right]\mathrm{e}^{-i\omega R/\beta} \end{aligned}$$

4.02.3 Moment Tensor Sources

The Green function for a point force is the fundamental solution of the equation of elastodynamics. However – except for a few rare exceptions – seismic sources are due to fast internal deformation in the Earth, for instance, faulting or fast phase changes on localized volumes inside the Earth. For a seismic source to be of internal origin, it has to have zero net force and zero net moment. It is not difficult to imagine seismic sources that satisfy these two conditions:

$$\begin{aligned} \sum \mathbf{f} &= 0 \\ \sum \mathbf{f} \times \mathbf{r} &= 0 \end{aligned} \qquad [18]$$

The simplest such sources are dipoles and quadrupoles. For instance, the so-called linear dipole is made of two identical point forces of strength f that act in opposite directions at two points separated by a very small distance h along the axes of the forces. The seismic moment of this linear dipole is $M_0 = f\,h$. Experimental observation has shown that linear dipoles of this sort are not the most frequent models of seismic sources and, furthermore, there does not seem to be any simple internal deformation mechanism that corresponds to a pure linear dipole. It is possible to combine three orthogonal linear dipoles in order to form a general seismic source; any dipolar seismic source can be simulated by adjusting the strength of these three dipoles. It is obvious, as we will show later, that these three dipoles represent the principal directions of a symmetrical tensor of rank two that we call the seismic moment tensor:

$$\mathbf{M} = \begin{bmatrix} M_{xx} & M_{xy} & M_{xz} \\ M_{xy} & M_{yy} & M_{yz} \\ M_{xz} & M_{yz} & M_{zz} \end{bmatrix}$$

This moment tensor has a structure that is identical to that of a stress tensor, but it is not of elastic origin as we shall see promptly.

What do the off-diagonal elements of the moment tensor represent? Let us consider a moment tensor such that all elements are zero except M_{xy}. This moment tensor represents a double couple, a pair of two couples of forces that turn in opposite directions. The first of these couples consists in two forces of direction $\mathbf{e}_x$ separated by a very small distance h in the direction y. The other couple consists in two forces of direction $\mathbf{e}_y$ with a small arm in the direction x. The moment of each of

these couples is M_{xy}, the first pair has positive moment, and the second has a negative one. The conditions of conservation of total force and moment [18] are satisfied so that this source model is fully acceptable from a mechanical point of view. In fact, as shown by Burridge and Knopoff (1964), the double couple is the natural representation of a fault. One of the pair of forces is aligned with the fault, the forces indicate the directions of slip, and the arm is in the direction of the fault normal.

4.02.3.1 Radiation from a Point Moment Tensor Source

Let us now use the Green functions obtained for a point force in order to calculate the radiation from a point moment tensor source located at point $\mathbf{x}_0$:

$$\mathbf{M}_0(\mathbf{r}, t) = \mathbf{M}_0(t)\delta(\mathbf{x} - \mathbf{x}_0) \qquad [19]$$

$\mathbf{M}_0$ is the moment tensor, a symmetrical tensor whose components are independent functions of time.

We consider one of the components of the moment tensor, for instance, M_{ij}. This represents two point forces of direction i separated by an infinitesimal distance h_j in the direction j. The radiation of each of the point forces is given by the Green function G_{ij} computed in [17]. The radiation from the M_{ij} moment is then simply

$$u_k(\mathbf{x}, t) = \sum_{ij} G_{ki,j}(\mathbf{x}, t|\mathbf{x}_0, t) * M_{ij}(t) \qquad [20]$$

The complete expression of the radiation from a point moment tensor source can then be obtained from [17]. We will be interested only on the far-field terms since the near field is too complex to discuss here. We get for the far-field waves

$$\begin{aligned} u_i^{\mathrm{P}}(R, t) &= \frac{1}{4\pi\rho\alpha^3}\frac{1}{R}\sum_{jk}\mathfrak{R}_{ijk}^{\mathrm{P}}\dot{M}_{jk}(t - R/\alpha) \\ u_i^{\mathrm{S}}(R, t) &= \frac{1}{4\pi\rho\beta^3}\frac{1}{R}\sum_{jk}\mathfrak{R}_{ijk}^{\mathrm{S}}\dot{M}_{jk}(t - R/\beta) \end{aligned} \qquad [21]$$

where $\mathfrak{R}_{ijk}^{\mathrm{P}} = R_i R_j R_k$ and $\mathfrak{R}_{ijk}^{\mathrm{S}} = (\delta_{ij} - R_i R_j)R_k$ are the radiation patterns of P- and S-waves, respectively. We observe in [21] that the far-field signal carried by both P- and S-waves is the time derivative of the seismic moment components, so that far-field seismic waves are proportional to the moment rate of the source.

Very often in seismology, it is assumed that the geometry of the source can be separated from its time variation, so that the moment tensor can be written in the simpler form

$$\mathbf{M}_0(t) = \mathbf{M}\, s(t)$$

where $\mathbf{M}$ is a time-invariant tensor that describes the geometry of the source and $s(t)$ is the time variation of the moment, the source time function determined by seismologists. Referring to Figure 2, we can now write a simpler form of [21]

$$u^{\mathrm{c}}(\mathbf{x}, t) = \frac{1}{4\pi\rho c^3}\frac{\mathfrak{R}^{\mathrm{c}}(\theta, \phi)}{R}\Omega(t - R/c) \qquad [22]$$

where R is the distance from the source to the observer. c stands for either α for P-waves or β for shear waves (SH and SV). For P-waves, u^c is the radial component; for S-waves, it is the appropriate transverse component for SH or SV waves. In

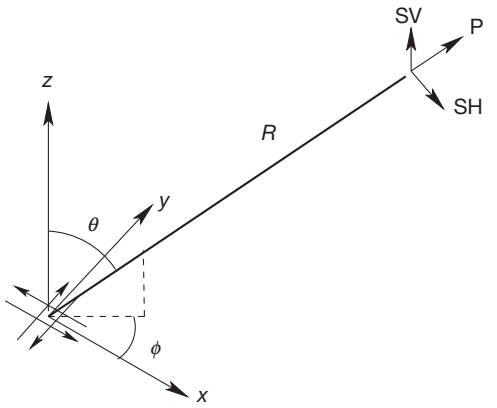


Figure 2 Radiation from a point double source. The source is at the origin and the observer at a position defined by the spherical coordinates R, θ, ϕ, distance, polar angle, and azimuth.

[22], we have introduced the standard notation $\Omega(t) = \dot{s}(t)$ for the source time function, the signal emitted by the source in the far field.

The term $\mathfrak{R}^{\mathrm{c}}(\theta, \phi)$ is the radiation pattern, a function of the takeoff angle of the ray at the source. Let (R, θ, ϕ) be the radius, colatitude, and azimuth of a system of spherical coordinates centered at the source, respectively. It is not difficult to show that the radiation pattern is given by

$$\mathfrak{R}^{\mathrm{c}}(\theta, \phi) = \mathbf{e}_{\mathrm{R}} \cdot \mathbf{M} \cdot \mathbf{e}_{\mathrm{R}} \qquad [23]$$

for P-waves, where $\mathbf{e}_{\mathrm{R}}$ is the radial unit vector at the source. Assuming that the z-axis at the source is vertical, so that θ is measured from that axis, S-waves are given by

$$\mathfrak{R}^{\mathrm{SV}}(\theta, \phi) = \mathbf{e}_\theta \cdot \mathbf{M} \cdot \mathbf{e}_{\mathrm{R}} \quad \text{and} \quad \mathfrak{R}^{\mathrm{SH}}(\theta, \phi) = \mathbf{e}_\phi \cdot \mathbf{M} \cdot \mathbf{e}_{\mathrm{R}} \qquad [24]$$

where $\mathbf{e}_\phi$ and $\mathbf{e}_\theta$ are unit vectors in spherical coordinates. Thus, the radiation patterns are the radial components of the moment tensor projected on spherical coordinates.

With minor changes to take into account smooth variations of elastic wave speeds in the Earth, these expressions are widely used to generate synthetic seismograms in the so-called far-field approximation. The main changes that are needed are the use of travel time $T_{\mathrm{c}}(r, r_{\mathrm{o}})$ instead of R/c in the waveform $\Omega(t - T_{\mathrm{c}})$ and a more accurate geometric spreading factor $g(\Delta, H)/a$ to replace $1/R$, where a is the radius of the Earth and $g(\Delta, H)$ is a tabulated function that depends on the angular distance Δ between the hypocenter and the observer and the source depth H. In most work with local earthquakes, the approximation [22] is frequently used with a simple correction for free surface response.

4.02.3.2 A More General View of Moment Tensors

What does a seismic moment represent? A number of mechanical interpretations are possible. In the previous sections, we introduced it as a simple mechanical model of double couples and linear dipoles. Other authors Eshelby (1956) and Backus and Mulcahy (1976) have explained them in terms of the distribution of inelastic stresses (sometimes called stress 'glut').

Let us first notice that a very general distribution of force that satisfies the two conditions [18] necessarily derives from a symmetrical seismic moment density of the form

$$\mathbf{f}(\mathbf{x}, t) = \nabla \cdot \mathbf{M}(\mathbf{x}, t) \qquad [25]$$

where $\mathbf{M}(\mathbf{x}, t)$ is the moment tensor density per unit volume. Gauss' theorem can be used to prove that such a force distribution has no net force nor moment. In many areas of applied mathematics, the seismic moment distribution is often termed a 'double layer potential.'

We can now use [25] in order to rewrite the elastodynamic eqn [1] as a system of first-order partial differential equations:

$$\begin{aligned} \rho \frac{\partial}{\partial t}\mathbf{v} &= \nabla \cdot \sigma \\ \frac{\partial}{\partial t}\sigma &= \lambda \nabla \cdot \mathbf{v}\mathbf{I} + \mu\left[(\nabla \mathbf{v}) + (\nabla \mathbf{v})^{\mathrm{T}}\right] + \dot{\mathbf{M}}_0 \end{aligned} \qquad [26]$$

where $\mathbf{v}$ is the particle velocity and σ is the corresponding elastic stress tensor. We observe that the moment tensor density source appears as an addition to the elastic stress rate $\dot{\sigma}$. This is probably the reason that Backus and Mulcahy (1976) adopted the term 'glut.' In many other areas of mechanics, the moment tensor is considered to represent the stresses produced by inelastic processes. A full theory of these stresses was proposed by Eshelby (1956). Incidentally, the equation of motion written in this form is the basis of some very successful numerical methods for the computation of seismic wave propagation see, for example, Madariaga (1976), Virieux (1986), and Madariaga et al. (1998).

We can get an even clearer view of the origin of the moment tensor density by considering it as defining an *inelastic* strain tensor ε^{I} defined implicitly by

$$(m_0)_{ij} = \lambda \delta_{ij} \varepsilon^{\mathrm{I}}_{kk} + 2\mu \varepsilon^{\mathrm{I}}_{ij} \qquad [27]$$

Many seismologists have tried to use ε_{I} in order to represent seismic sources. Sometimes termed 'potency' (Ben Menahem and Singh, 1981), the inelastic strain has not been widely adopted even if it is a more natural way of introducing seismic source in bimaterial interfaces and other heterogeneous media. For a recent discussion, see Ampuero and Dahlen (2005).

The meaning of ε^{I} can be clarified by reference to **Figure 3**. Let us make the following 'gedanken'experiment. Let us cut an infinitesimal volume V from the source region. Next, we let it undergo some inelastic strain ε^{I}, for instance, a shear strain due to the development of internal dislocations as shown in the figure. Let us now apply stresses on the borders of the internally deformed volume V so as to bring it back to its original shape. If the elastic constants of the internally deformed volume V have not changed, the stresses needed to bring V back to its original shape are exactly given by the moment tensor components defined in [27]. This is the definition of seismic moment tensor: It is the stress produced by the inelastic deformation of a body that is elastic everywhere. It should be clear that the moment tensor is not the same thing as the stress tensor acting on the fault zone. The latter includes the elastic response to the introduction of internal stresses as shown in the last row of **Figure 3**. The difference between the initial stresses before the internal deformation and those that prevail after the deformed body has been reinserted in the elastic medium is the stress change (or stress drop as originally introduced in seismology in the late 1960s). If the internal strain is produced in the sense of reducing applied stress – and reducing internal strain energy – then stresses inside the source will decrease in a certain average sense. It must be understood, however, that a source of internal origin – like faulting – can only redistribute internal stresses. During faulting, stresses reduce in the immediate vicinity of slip zones but increase almost everywhere else.

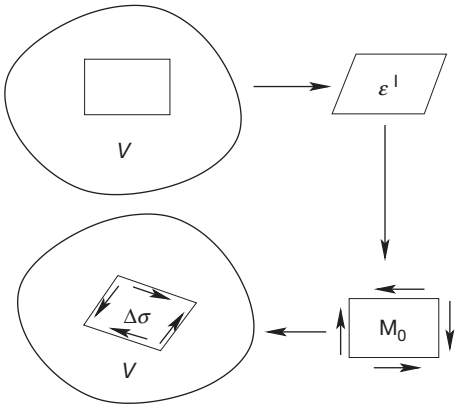


Figure 3 Inelastic stresses or stress glut at the origin of the concept of seismic moment tensor. We consider a small rectangular zone that undergoes a spontaneous internal deformation ϵ^{I} (top row). The elastic stresses needed to bring it back to a rectangular shape are the moment tensor or stress glut (bottom row right). Once stresses are relaxed by interaction with the surrounding elastic medium, the stress change is $\Delta\sigma$ (bottom left).

4.02.3.3 Moment Tensor Equivalent of a Fault

For a point moment tensor of type [27], we can write

$$(M_0)_{ij} = \left(\lambda \delta_{ij} \varepsilon^{\mathrm{I}}_{kk} + 2\mu \varepsilon^{\mathrm{I}}_{ij}\right) V \delta(\mathbf{x} - \mathbf{x}_0) \qquad [28]$$

where V is the elementary source volume on which acts the source. Let us now consider that the source is a very thin cylinder of surface S, thickness h, volume $V = Sh$, and unit normal $\mathbf{n}$. Letting the thickness of the cylinder tend to zero, the mean inelastic strain inside the volume V can be computed as follows:

$$\lim_{h \to 0} \varepsilon^{\mathrm{I}}_{ij} h = \frac{1}{2}\left[\Delta u_i n_j + \Delta u_j n_i\right] \qquad [29]$$

where Δu is the displacement discontinuity (or simply the *slip*) across the fault volume. The seismic moment for the flat fault is then

$$(M_0)_{ij} = \left[\lambda \delta_{ij} \Delta u_k n_k + \mu\left(\Delta u_i n_j + \Delta u_j n_i\right)\right] S \qquad [30]$$

so that the seismic moment can be defined for a fault as the product of an elastic constant with the displacement discontinuity and the source area. Actually, this is the way the seismic moment was originally determined by Burridge and Knopoff (1964). If the slip discontinuity is written in terms of a direction of slip v and a scalar slip D, $\Delta u_i = D v_i$, we get

$$(M_0)_{ij} = \delta_{ij} v_k n_k \lambda D S + \left(v_i n_j + v_j n_i\right) \mu D S \qquad [31]$$

Most seismic sources do not produce normal displacement discontinuities (fault opening) so that $v \cdot \mathbf{n} = 0$ and the first

term in [30] is equal to zero. In that case, the seismic moment tensor can be written as the product of a tensor with the scalar seismic moment $M_0 = \mu DS$:

$$(M_0)_{ij} = (v_i n_j + v_j n_i)\mu DS \qquad [32]$$

The first practical determination of the scalar seismic moment is due to Aki (1966), who estimated M_0 from seismic data recorded after the Niigata earthquake of 1966 in Japan. Determination of seismic moment has become the standard way in which earthquakes are measured. All sorts of seismological, geodetic, and geologic techniques have been used to determine M_0. A worldwide catalog of seismic moment tensors was made available online by Harvard University (Dziewonski and Woodhouse, 1983). Initially, moments were determined by for the limited form [32], but nowadays, the full set of six components of the moment tensor is regularly computed.

Let us remark that the restricted form of the moment tensor [32] reduces the number of independent parameters of the moment tensor from 6 to 4. Very often, seismologists use the simple fault model of the source moment tensor. In practice, the fault is parameterized by the seismic moment plus the three Euler angles of the fault plane. Following the convention adopted by Aki and Richards (2002), these angles are defined as δ the dip of the fault, ϕ the strike of the fault with respect to the north, and λ the rake of the fault, which is the angle of the slip vector with respect to the horizontal.

4.02.3.4 Eigenvalues and Eigenvectors of the Moment Tensor

Since the moment tensor is a symmetrical tensor of order 3, it has three orthogonal eigenvectors with real eigenvalues, just like any stress tensor. These eigenvalues and eigenvectors are the three solutions of

$$\mathbf{M}_0\mathbf{v} = m\mathbf{v}$$

Let the eigenvalues and eigenvector be

$$m_i, \quad \mathbf{v}_i \qquad [33]$$

Then, the moment tensor can be rewritten as

$$\mathbf{M}_0 = \sum_i m_i \mathbf{v}_i^{\mathrm{T}} \mathbf{v}_i \qquad [34]$$

Each eigenvalue–eigenvector pair represents a linear dipole. The eigenvalue represents the moment of the dipole. From extensive studies of moment tensor sources, it appears that many seismic sources are very well represented by an almost pure double-couple model with $m_1 = -m_3$ and $m_2 \cong 0$.

A great effort for calculating moment tensors for deeper sources has been made by several authors in recent years. It appears that the non-double couple part is larger for these sources but that it does not dominate the radiation. For deep sources, Knopoff and Randall (1970) proposed the so-called compensated linear vector dipole (CLVD). This is a simple linear dipole from which we subtract the volumetric part so that $m_1 + m_2 + m_3 = 0$. Thus, a CLVD is a source model where $m_2 = m_3 = -1/2 m_1$. Radiation from a CLVD is very different from that from a double-couple model, and many seismologists have tried to separate a double couple from CLVD components from the moment tensor. In fact, moment tensors are better represented by their eigenvalues; separation into a fault and a CLVD part is generally ambiguous.

Seismic moments are measured in units of Nm. Small earthquakes that produce no damage have seismic moments less than 10^{12} Nm, while the largest subduction events, like those of Chile in 1960, Alaska in 1964, and Sumatra in 2004, have moments of the order of 10^{22}–10^{23} Nm. Large destructive events like Izmit, Turkey, 1999; Chichi, Taiwan, 1999; and Landers, California, 1992 have moments of the order of 10^{20} Nm.

Since the late 1930s, it became commonplace to measure earthquakes by their magnitude, a logarithmic measure of the total energy radiated by the earthquake. Methods for measuring radiated energy were developed by Gutenberg and Richter using well-calibrated seismic stations. At the time, the general properties of the radiated spectrum were not known and the concept of moment tensor had not yet been developed. Since at present time earthquakes are systematically measured using seismic moments, it has become standard to use the following empirical relation defined by Kanamori (1977) and Hanks and Kanamori (1979) to convert moment tensors into a magnitude scale:

$$\log_{10} M_0 \quad (\text{in Nm}) \quad = 1.5 M_w + 9 \qquad [35]$$

Magnitudes are easier to retain and have a clearer meaning for the general public than the more difficult concept of moment tensor.

4.02.3.5 Seismic Radiation from Moment Tensor Sources in the Spectral Domain

In actual applications, the near-field signals radiated by earthquakes may become quite complex because of multipathing, scattering, etc., so that the actually observed seismogram, say, $u(t)$, resembles the source time function $\Omega(t)$ only at long periods. It is usually verified that complexities in the wave propagation affect much less the spectral amplitudes in the Fourier transformed domain. Radiation from a simple point moment tensor source can be obtained from [22] by straightforward Fourier transformation:

$$\tilde{u}_c(\mathbf{x}, \omega) = \frac{1}{4\pi\rho c^3} \frac{\Re^c(\theta_0, \phi_0)}{R} \tilde{\Omega}(\omega) e^{-i\omega R/c} \qquad [36]$$

where $\tilde{\Omega}(\omega)$ is the Fourier transform of the source time function $\Omega(t)$. A well-known property of Fourier transform is that the low-frequency limit of the transform is the integral of the source time function, that is,

$$\lim_{\omega \to 0} \tilde{\Omega}(\omega) = \int_0^\infty \dot{M}_0(t) \mathrm{d}t = M_0$$

So that in fact, the low-frequency limit of the transform of the displacement yields the total moment of the source. Unfortunately, the same notation is used to designate the total moment release by an earthquake – M_0 – and the time-dependent moment $M_0(t)$.

From the observation of many earthquake spectra, and from the scaling of moment with earthquake size, Aki (1967) and Brune (1970) concluded that the seismic spectra decayed as ω^{-2} at high frequencies. Although, in general, spectra are more complex for individual earthquakes, a simple source model can be written as follows:

$$\Omega(\omega) = \frac{M_0}{1 + (\omega/\omega_0)^2} \quad [37]$$

where ω_0 is the so-called corner frequency. In this simple 'omega-squared model,' seismic sources are characterized by only two independent scalar parameters: the seismic moment M_0 and the corner frequency ω_0. Not all earthquakes have displacement spectra as simple as [37], but the omega-squared model is a convenient starting point for understanding seismic radiation.

From [37], it is possible to compute the spectra of ground velocity $\dot{\Omega}(\omega) = i\omega\Omega(\omega)$. Ground velocity spectra have a peak situated roughly at the corner frequency ω_0. In actual earthquake spectra, this peak is usually broadened and contains oscillations and secondary peaks; at higher frequencies, attenuation, propagation scattering, and source effects reduce the velocity spectrum.

Finally, by an additional differentiation, we get the acceleration spectra $\ddot{\Omega}(\omega) = -\omega^2\Omega(\omega)$. This spectrum has an obvious problem; it predicts that ground acceleration is flat for arbitrarily high frequencies. The acceleration spectrum usually decays after a high-frequency corner identified as f_{max}. The origin of this high-frequency cutoff was a subject of discussion in the 1990s, which was settled by the implicit agreement that f_{max} reflects the dissipation of high-frequency waves due to propagation in a strongly scattering medium, like the crust and near-surface sediments.

It is interesting to observe that [37] is the Fourier transform of

$$\Omega(t) = \frac{M_0\omega_0}{2} e^{-|\omega_0 t|} \quad [38]$$

This is a noncausal strictly positive function, is symmetrical about the origin, and has an approximate width of $1/\omega_0$. By definition, the integral of the function is exactly equal to M_0. Even if this function is noncausal, it shows that $1/\omega_0$ controls the width or duration of the seismic signal. At high frequencies, the function behaves like ω^{-2}. This is due to the slope discontinuity of [38] at the origin.

We can also interpret [37] as the absolute spectral amplitude of a causal function. There are many such functions, one of them – proposed by Brune (1970) – is

$$\Omega(t) = M_0\omega_0^2 t\, e^{-\omega_0 t} H(t) \quad [39]$$

As for [38], the width of the function is roughly $1/\omega_0$ and the high frequencies are due to the slope break of $\Omega(t)$ at the origin. This slope break has the same amplitude as that of [38].

4.02.3.6 Seismic Energy Radiated by Point Moment Tensor Sources

In Section 4.02.2.4, we discussed the general energy balance for a seismic source embedded in an elastic medium. The energy flow for a moment tensor source can be derived from expression [14], where we replace the force density by its expression in terms of a moment density [25]. After a few small changes and integration by parts, we get

$$E_s = -\Delta U + \int_0^\infty dt \int_V M_{ij}\dot{\varepsilon}_{ij} dV \quad [40]$$

Seismic radiation is equal to the reduction in internal strain energy plus the work of the moment tensor against the elastic strain rate at the source.

As we have already discussed for a point force, at any position sufficiently far from the source, energy flow per unit solid angle is proportional to the square of local velocity [15]. Using the far-field approximation [36], we can follow the same steps as in [16] to express the radiated energy in terms of the seismic source time function as

$$E_c = \frac{1}{4\pi\rho c^5}\langle \Re^c \rangle_2 \int_0^\infty \dot{\Omega}^2(t) dt$$

where c stands again for P- or S-waves and, as in [18], $\langle \Re^i \rangle_2 = 1/(4\pi)\int_\Omega (\Re^i)^2 d\Omega$ is the mean square radiation pattern. We can now use Parseval's theorem

$$\int_0^\infty \dot{\Omega}(t)^2 dt = \frac{1}{\pi}\int_0^\infty \omega^2 \widetilde{\Omega}(\omega)^2 d\omega$$

in order to express the radiated energy in terms of the seismic spectrum as

$$E_c = \frac{1}{4\pi^2\rho c^5}\langle \Re^c \rangle_2 \int_0^\infty \omega^2\Omega^2(\omega) d\omega \quad [41]$$

For Brune's spectrum [37], the integral is

$$\int_0^\infty \omega^2\Omega^2(\omega) d\omega = \frac{\pi}{2} M_0^2\omega_0^3$$

so that radiated energy is proportional to the square of moment. We can finally write

$$\frac{E_c}{M_0} = \frac{1}{16\pi\rho c^5}\langle \Re^c \rangle^2 M_0\omega_0^3 \quad [42]$$

This nondimensional relation makes no assumptions about the rupture process at the source except that the spectrum is of the form [37]. Noting that in the Earth, α is roughly $\sqrt{3}\beta$ so that $\alpha^5 \cong 16\beta^5$, the amount of energy carried by the S-waves emitted by a point moment tensor is close to 25 times that carried by P-waves, if the source spectrum is the same for P- and S-waves. For finite sources, the partition of energy into P- and S-waves depends on the details of the rupture process.

Since the energy flow e_c can usually be determined in only a few directions, (θ, ϕ) of the focal sphere, the energy-moment ratio [42] can only be estimated, never computed very precisely. This problem still persists; in spite of the deployment of increasingly denser instrumental networks, there will always be large areas of the focal sphere that remain out of the domain of seismic observations because the waves in those directions are refracted away from the station networks, energy is dissipated due to long trajectories, etc.

4.02.3.7 More Realistic Radiation Models

In reality, earthquakes occur in a complex medium that is usually scattering and dissipative. Seismic waves become diffracted and reflected and in general suffer multipathing in those structures. Accurate seismic modeling would require perfect knowledge of those structures. It is well known and understood that those complexities dominate signals in certain frequency bands. For this reason, the simple model presented

here can be used to understand many features of earthquakes, and the more sophisticated approaches that attempt to model every detail of the waveform are reserved only for more advanced studies. Here, like in many other areas of geophysics, a balance between simplicity and concepts must be kept against numerical complexity that may not always be warranted by lack of knowledge of the details of the structures. If the simple approach were not possible, then many standard methods to study earthquakes would be impossible to use. For instance, source mechanism, the determination of the fault angles δ, ϕ, and λ, would be impossible. These essential parameters are determined by back projection of the displacement directions from the observer to a virtual unit sphere around the point source.

A good balance between simple, but robust concepts, and the sophisticated reproduction of the complex details of real wave propagation is a permanent challenge for seismologists. Nowadays, numerical techniques become more and more common. Our simple models detailed earlier are not to be easily neglected; in any case, they should always serve as test models for fully numerical methods.

4.02.4 Finite Source Models

The point source model we just discussed provides a simple approach to the simulation of seismic radiation. It is probably quite sufficient for the purpose of modeling small sources situated sufficiently far from the observer so that the source looks like a single point source. Details of the rupture process are then hidden inside the moment tensor source time function $M_0(t)$. For larger earthquakes, and especially for earthquakes observed at distances close to the source, the point source model is not sufficient, and one has to take into account the geometry of the source and the propagation of rupture across the fault. Although the first finite models of the source are quite ancient, their widespread use to model earthquakes is relatively recent and has been more extensively developed as the need to understand rupture in detail has been more pressing. The first models of a finite fault were developed simultaneously by Maruyama (1963) and Burridge and Knopoff (1964) in the general case, by Ben Menahem (1961, 1962) for surface and body waves, and by Haskell (1964, 1966) who provided a very simple solution for the far field of a rectangular fault. Haskell's model became the *de facto* earthquake fault model in the late 1960s and early 1970s and was used to model many earthquakes. In the following, we review the available finite source models, focusing on the two main models: the rectangular fault and the circular fault.

4.02.4.1 The Kinematic Dislocation Model

In spite of much recent progress in understanding the dynamics of earthquake ruptures, the most widely used models for interpreting seismic radiation are the so-called dislocation models. In these models, the earthquake is simulated as the kinematic spreading of a displacement discontinuity (slip or dislocation in seismological usage) along a fault plane. As long as the thickness of the fault zone H is negligible with respect to the other length scales of the fault (width W and length L), the fault may be idealized as a surface of displacement discontinuity or slip. Slip is very often called dislocation by seismologists, although this is not the same as the concept of a dislocation in solid mechanics.

In its most general version, slip as a function of time and position in a dislocation model is completely arbitrary, and rupture propagation may be as general as wanted. In this version, the dislocation model is an appropriate description of an earthquake as the propagation of a slip episode on a fault plane. It must be remarked, however, that not all slip distributions are physically acceptable. Madariaga (1978) showed that Haskell's model, by far the most used dislocation model, presents unacceptable features like interpenetration of matter that make it very difficult to use at high frequencies without important modifications. The most serious problem with Haskell's model is that strain energy release is infinite, so that this model is not useful for the study of seismic energy balance. For this reason, dislocation models must be considered as an intermediate step in the formulation of a physically acceptable description of rupture but examined critically when converted into dynamic models. From this perspective, dislocation models are very useful step in the inversion of near-field accelerograms (see, e.g., Wald and Heaton, 1994).

A finite source model can be described as a distribution of moment tensor sources. Since we are interested in radiation from faults, we use the approximation [32] for the moment of a fault element. Each of these elementary sources produces a seismic radiation that can be computed using the Green function [17]. The total displacement seismogram observed at an arbitrary position $\mathbf{x}$ is the sum

$$u_i(\mathbf{x}, t) = \int_0^t \int_S \mu(\mathbf{x}_0) \Delta u_j(\mathbf{x}_0, \tau) G_{ij,k}(\mathbf{x}, t|\mathbf{x}_0, \tau) n_k(\mathbf{x}_0) \mathrm{d}^2\mathbf{x}_0 \mathrm{d}\tau \quad [43]$$

where $\Delta u(\mathbf{x}_0, t)$ is the slip across the fault of surface S as a function of position on the fault ($\mathbf{x}_0$) and time t. $\mathbf{n}$ is the normal to the fault and $\mathbf{G}(x, t)$ is the elastodynamic Green tensor that may be computed using simple layered models of the crustal structure or more complex finite difference simulations.

In a first, simple analysis, we can use the far-field approximation [22] that is often used to generate synthetic seismograms far from the source (see, e.g., Kikuchi and Kanamori, 1982, 1991). Inserting [22] into [43] and after some simplification, we get

$$u^c(\mathbf{x}, t) = \frac{1}{4\pi\rho c^3} \int_0^t \int_S \frac{\mathfrak{R}^c_{ijk}(\theta, \phi)}{R} \mu \Delta \dot{u}_j \left[\mathbf{x}_0, t - \tau - \frac{R(\mathbf{x} - \mathbf{x}_0)}{c}\right] n_k \mathrm{d}^2\mathbf{x}_0 \mathrm{d}\tau \quad [44]$$

where $R(\mathbf{x} - \mathbf{x}_0)$ is the distance between the observer and a source point located at $\mathbf{x}_0$. In almost all applications, the reference point is the hypocenter, the point where rupture initiates.

In [44], both the radiation pattern $\mathfrak{R}^c$ and the geometric decay $1/R$ change with position on the fault. In the far field, according to ray theory, we can make the approximation that only travel time changes are important so that we can approximate the integral [44] assuming that both radiation pattern and geometric spreading do not change significantly over the fault. In the far field, we can also make the so-called Fraunhofer approximation:

$$R(\mathbf{x}-\mathbf{x}_0) \cong R(\mathbf{x}-\mathbf{x}_H) - \mathbf{e}_r \cdot (\mathbf{x}_0 - \mathbf{x}_H)$$

where $\mathbf{x}_H$ is a reference point on the fault, usually the hypocenter, and $\mathbf{e}_r$ is the unit vector in the radial direction from the reference point to the observer. With these approximations, far-field radiation from a finite source is again given by the generic expression [22] where the source time function Ω is replaced by

$$\Omega(t, \theta, \phi) = \mu \int_0^t \int_S \Delta \dot{u}_j \left[\xi_1, \xi_2, t - \tau + \frac{\mathbf{e}_r \cdot \xi}{c} \right] d\xi_1 d\xi_2 d\tau \qquad [45]$$

where ξ is a vector of component (ξ_1, ξ_2) that measures position on the fault with respect to the hypocenter $\mathbf{x}_H$. The main difference between a point and a finite source as observed from the far field is that in the finite case, the source time function Ω depends on the direction of radiation (θ, ϕ) through the term $\mathbf{e}_r \cdot \xi$. This directivity of seismic radiation can be very large when ruptures propagate at high subshear or intersonic speeds (**Chapter 4.09**).

The source time function expression [45] is linear in slip rate amplitude but very nonlinear with respect to rupture propagation, which is implicit in the time dependence of $\Delta \dot{u}$. For this reason, in most inversions, the kinematics of the rupture process (position of rupture front as a function of time) is simplified. The most common assumption is to assume that rupture propagates at constant speed away from the hypocenter. Different approaches have been proposed in the literature in order to approximately invert for variations in rupture speed about the assumed constant rupture velocity (see, e.g., Cotton and Campillo, 1995; Hartzell and Heaton, 1983; Wald and Heaton, 1994; **Chapter 4.09**).

4.02.4.1.1 Haskell's rectangular fault model

One of the most widely used dislocation model was introduced by Haskell (1964, 1966). In this model, shown in **Figure 4**, a uniform displacement discontinuity spreads at constant rupture velocity inside a rectangular-shaped fault. Although from a mechanical point of view this model violates simple principles of continuum mechanics, like continuity of matter, it is a very useful first approximation to a seismic source. At low frequencies, or wavelengths much longer than the size of the fault, this model is a reasonable approximation to a simple seismic rupture propagating along a strike-slip fault.

In Haskell's model, at time $t = 0$, a line of dislocation of width W appears suddenly and propagates along the fault at a constant rupture velocity until a region of length L of the fault has been broken. As the dislocation moves, it leaves behind a zone of constant slip D. Assuming that the fault lies on a plane of coordinates (ξ_1, ξ_2), the slip function can be written as (see also **Figure 4**)

$$\Delta \dot{u}_1(\xi_1, \xi_2, t) = D\dot{s}(t - \xi_1/v_r) H(\xi_1) H(L - \xi_1) \quad \text{for} \quad -W/2 < \xi_2 < W/2 \qquad [46]$$

where $\dot{s}(t)$ is the slip rate time function that in the simplest version of Haskell's model is invariant with position on the fault. The most important feature of this model is the propagation of rupture implicit in the time delay of rupture. $\xi_1/v_r \cdot v_r$ is the rupture velocity, the speed with which the slip front propagates along the fault in the ξ_1 direction. An obvious unphysical feature of this model is that rupture appears instantaneously in the ξ_2 direction, this is of course impossible for a spontaneous seismic rupture. The other inadmissible feature of Haskell's model is the fact that on its borders, slip suddenly jumps from the average slip D to zero. This violates material continuity so that the most basic equation of motion (1) is no longer valid near the edges of the fault. In spite of these two obvious shortcomings, Haskell's model gives a simple, first-order approximation to seismic slip, fault finiteness at a finite rupture speed. The seismic moment of Haskell's model is easy to compute, the fault area is $L \times W$, and slip D is constant over the fault, so that the seismic moment is $M_0 = \mu DLW$. Using the far-field approximation, we can compute the source time function for Haskell's model:

$$\Omega_H(t, \theta, \phi) = \mu \int_{-W/2}^{W/2} d\xi_2 \int_0^L D\dot{s} \left[t - \frac{\xi_1}{v_r} + \frac{\xi_1}{c} \cos\phi \sin\theta + \frac{\xi_2}{c} \sin\phi \sin\theta \right] d\xi_1 \qquad [47]$$

where we used the index H to indicate that this is Haskell's model. The two integrals can be evaluated very easily for an observer situated along the axis of the fault, that is, when $\phi = 0$.

Integrating, we get

$$\Omega_H(\theta, 0, t) = M_0 \frac{1}{T_M} \int_0^{\min(t, T_M)} \dot{s}(t - \tau) d\tau \qquad [48]$$

where $T_M = L/c(1 - v_r/c \sin\theta)$. Thus, the far-field signal is a simple integral over time of the source slip rate function. In other directions, the source time function Ω_H is more complex, but it can be easily computed by the method of isochrones that we will explain shortly. Radiation from Haskell's model shows two very fundamental properties of seismic radiation: finite duration given by T_M and directivity, that is, the duration and amplitude of seismic waves depend on the radiation angles θ and ϕ.

A similar computation in the frequency domain was made by Haskell (1966) for the particular direction $\phi = 0$. In our notation, the result is

$$\tilde{\Omega}_H(\theta, 0, \omega) = M_0 \mathrm{sinc}(\omega T_M/2) e^{-i\omega T_M/2} \tilde{\dot{s}}(\omega) \qquad [49]$$

where $\mathrm{sinc}(x) = sin(x)/x$.

It is often assumed that the slip rate time function $\dot{s}(t)$ is a boxcar function of amplitude $1/\tau_r$ and duration τ_r, the rise time. In that case, the spectrum $\tilde{\Omega}(\omega)$ becomes

$$\tilde{\Omega}_H(\theta, 0, \omega) = M_0 \mathrm{sinc}(\omega T_M/2) \mathrm{sinc}(\omega \tau_r/2) e^{-i\omega(T_M + \tau_r)/2} \qquad [50]$$

or in the time domain

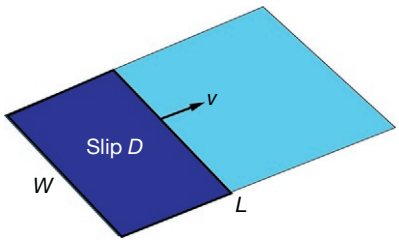


Figure 4 Haskell's kinematic model, one of the simplest possible earthquake models. The fault has a rectangular shape, and a linear rupture front propagates from one end of the fault to the other at constant rupture speed v. Slip in the broken part of the fault is uniform and equal to D.

$$\Omega_H(\theta, 0, t) = M_0 \text{boxcar}[t, T_M] * \text{boxcar}[t, \tau_r]$$

where the star means convolution and boxcar is a function of unit area that is zero everywhere except that in the time interval from 0 to τ_r where it is equal to $1/\tau_r$. Thus, Ω_H is a simple trapezoidal pulse of area M_0 and duration $T_d = T_M + \tau_r$. This surprisingly simple source time function matches the ω-squared model for the far-field spectrum since Ω_H is flat at low frequencies and decays like ω^{-2} at high frequencies. The spectrum has two corners associated with the pulse duration T_M and the other with rise time τ_r. This is unfortunately only valid for radiation along the plane $\phi = 0$ or $\phi = \pi$. In other directions with $\phi \neq 0$, radiation is more complex and the high-frequency decay is of order ω^{-3}, faster than in classical Brune's model.

In spite of some obvious mechanical shortcomings, Haskell's model captures some of the most important features of an earthquake and has been extensively used to invert for seismic source parameters in both the near field and far field from seismic and geodetic data. The complete seismic radiation for Haskell's model was computed by Madariaga (1978).

4.02.4.2 The Circular Fault Model

The other simple source model that has been widely used in earthquake source seismology is a circular crack model. This model was introduced by several authors including Savage (1966), Brune (1970), and Keylis-Borok (1959) to quantify a simple source model that was mechanically acceptable and to relate slip on a fault to stress changes. As we have already mentioned, dislocations models like Haskell's produce nonintegrable stress changes due to the violation of material continuity at the edges of the fault. The circular crack model was a natural approach to model earthquakes avoiding nonphysical singularities. In the present section, we will examine the circular crack model from a kinematic point of view.

4.02.4.2.1 Kostrov's self-similar circular crack

The simplest possible crack model is that of a circular rupture that starts from a point and then spreads self-similarly at constant rupture speed v_r without ever stopping. Slip on this fault is driven by stress drop inside the fault. The solution of this problem is somewhat difficult to obtain because it requires very advanced use of self-similar solutions to the wave equation and its complete solution for displacements and stresses must be computed using the Cagniard–de Hoop method (Richards, 1976). Fortunately, the solution for slip across the fault found by Kostrov (1964) is surprisingly simple:

$$\Delta u_x(r, t) = C(v_r)\frac{\Delta\sigma}{\mu}\sqrt{v_r^2 t^2 - r^2} \qquad [51]$$

where r is the radius in a cylindrical coordinate system centered on the point of rupture initiation (see **Figure 5**). $V_r t$ is the instantaneous radius of the rupture at time t. $\Delta\sigma$ is the constant stress drop inside the rupture zone, μ is the elastic rigidity, and $C(v_r)$ is a very slowly varying function of the rupture velocity. For most practical purposes, $C \approx 1$. As shown by Kostrov (1964), inside the fault, the stress change produced by the slip function [51] is constant and equal to $\Delta\sigma$. This simple solution provides a very basic result that is one of the most important properties of circular cracks. Slip in the fault scales with the ratio of stress drop over rigidity times the instantaneous radius of the fault. As rupture radius increases, all the displacements around the fault scale with the size of the rupture zone.

The circular self-similar rupture model produces far-field seismic radiation with a very peculiar signature. Inserting the slip function into the expression for far-field radiation [45], we get

$$\Omega_K(t, \theta) = A(v_r, \theta) t^2 H(t)$$

where we used an index K to indicate Kostrov's model. The amplitude coefficient A is

$$A(v_r, \theta) = C(v_r)\frac{2\pi}{\left(1 - v_r^2/c^2 \sin^2\theta\right)^2}\Delta\sigma v_r^3$$

(see Boatwright, 1980; Richards, 1976). Thus, the initial rise of the far-field source time function is proportional to t^2 for Kostrov's model. The rate of growth is affected by a *directivity*

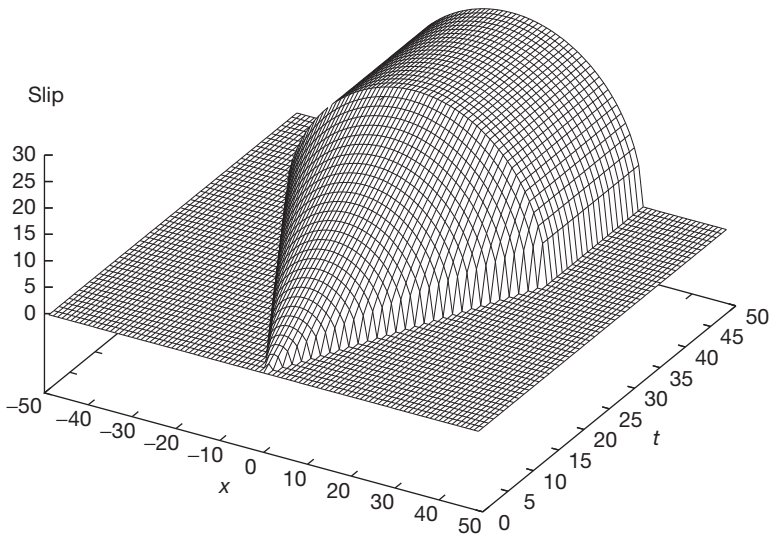


Figure 5 Slip distribution as a function of time on Sato and Hirasawa (1973) circular dislocation model.

factor in the denominator $(1-v_r^2/c^2\sin\theta)^2$ that increases with the polar angle θ and is maximum for $\theta=\pi/2$.

4.02.4.2.2 The kinematic circular source model of Sato and Hirasawa

Simple Kostrov's self-similar crack is not a good seismic source model for two reasons: (1) rupture never stops so that the seismic waves emitted by this source increase like t^2 without limit and (2) it does not explain the high-frequency radiation from seismic sources. Sato and Hirasawa (1973) proposed a modification of Kostrov's model that retained its initial rupture behavior [59] but added the stopping of rupture. They assumed that the Kostrov-like growth of the fault was suddenly stopped across the fault when rupture reached a final radius a (see **Figure 5**). In mathematical terms, the slip function is

$$\begin{aligned}\Delta u_x(r,t) &= C(v_r)\frac{\Delta\sigma}{\mu}\sqrt{v_r^2t^2-r^2}\,\mathrm{H}(v_rt-r) \quad \text{for} \quad t<a/v_r \\ &= C(v_r)\frac{\Delta\sigma}{\mu}\sqrt{a^2-r^2}\,\mathrm{H}(a-r) \quad \text{for} \quad t>a/v_r\end{aligned} \qquad [52]$$

Thus, at $t=a/v_r$, the slip on the fault becomes frozen and no motion occurs thereafter. This mode of healing is noncausal, but the solution is mechanically acceptable because slip near the borders of the fault always tapers like a square root of the distance to the fault tip. Using the far-field radiation approximation [45], Sato and Hirasawa found that the source time function for this model could be computed exactly

$$\Omega_{SH}(t,\theta) = C(v_r)\frac{2\pi}{\left(1-v_r^2/c^2\sin^2\theta\right)^2}\Delta\sigma v_r^3t^2 \qquad [53]$$

for $t<L/v_r(1-v_r/c\sin\theta)$ where θ is the polar angle of the observer. As should have been expected, the initial rise of the radiated field is the same as in Kostrov's model, the initial phase of the source time function increases very fast like t^2. After the rupture stops, the radiated field is

$$\Omega_{SH}(t,\theta) = C(v_r)\frac{\pi}{2}\frac{1}{v_r/c\sin\theta}\left[1-\frac{v_r^2t^2}{a^2(1+v_r/c\sin\theta)^2}\right]\Delta\sigma a^2v_r \qquad [54]$$

for times between $t_{s1}=a/v_r(1-v_r/c\sin\theta)$ and $t_{s2}=a/v_r(1+v_r/c\sin\theta)$. Radiation from the stopping process is spread in the time interval between the two stopping phases emitted from the closest (t_{s1}) and the farthest (t_{s2}) points of the fault. For times greater than t_{s2}, Ω returns to zero. The waves radiated by circular crack present both *directivity* in the second term and strong *focusing* due to the inverse sin θ term. Because of this term, the radiated field Ω is infinite along the axis of the fault.

It is also possible to compute the spectrum of the far-field signal ([53] and [54]) analytically. This was done by Sato and Hirasawa (1973). The important feature of the spectrum is that it is dominated by the stopping phases at times t_{s1} and t_{s2}. The stopping phases are both associated with a slope discontinuity of the source time function. As we already discussed for Brune's model, slope discontinuities produce ω^{-2} high-frequency asymptotes. This simple model explains one of the most universal features of seismic sources: The high frequencies radiated by seismic sources are dominated by stopping phases not by the energy radiated from the initiation of seismic rupture (Savage, 1966). These stopping phases are ubiquitous in dynamic models of faulting.

4.02.4.3 Generalization of Kinematic Models and the Isochrone Method

A simple yet powerful method for understanding the general properties of seismic radiation from classical dislocation models was proposed by Bernard and Madariaga (1984) and Spudich and Frazer (1984). The method was recently extended to study radiation from supershear ruptures by Bernard and Baumont (2005). The idea is that since most of the energy radiated from the fault comes from the rupture front, it should be possible to find where energy is coming from at a given station and at a given time. Bernard and Madariaga (1984) originally derived the isochrone method by inserting the ray-theoretical expression [22] into the representation theorem, a technique that is applicable not only in the far field but also in the immediate vicinity of the fault at high frequencies. Here, for the purpose of simplicity, we derive isochrones only in the far field. For that purpose, we study the far-field source time function for a finite fault derived in [45]. We assume that the slip rate distribution has the general form

$$\Delta\dot{u}_i(\xi_1,\xi_2,t) = D_i(t-\tau(\xi_1,\xi_2)) = D_i(t)*\delta(t-\tau(\xi_1,\xi_2)) \qquad [55]$$

where $\tau(\xi_1,\xi_2)$ is the rupture delay time at a point of coordinates ξ_1,ξ_2 on the fault. This is the time that it takes for rupture to arrive at that point. The star indicates time-domain convolution. We rewrote [55] as a convolution in order to distinguish between the slip time function $D(t)$ and its propagation along the fault described by the argument to the delta function. While we assume here that $D(t)$ is strictly the same everywhere on the fault, in the spirit of ray theory, our result remains valid if $D(\xi_1,\xi_2,t)$ is a slowly variable function of position on the fault. Inserting the slip rate field [55] in the source time function [45], we get

$$\Omega(t,\theta,\phi) = \mu D_i(t)*\int_0^t\int_{S_0}\delta\left[t-\tau(\xi_1,\xi_2)-\frac{\mathbf{e}\cdot\mathbf{x}_0}{c}\right]\mathrm{d}^2\mathbf{x}_0\mathrm{d}\tau \qquad [56]$$

where the star indicates time-domain convolution. Using the sifting property of the delta function, the integral over the fault surface S_0 reduces to an integral over a line defined implicitly by

$$t = \tau(\xi_1,\xi_2)+\frac{\mathbf{e}\cdot\mathbf{x}_0}{c} \qquad [57]$$

the solutions of this equation define one or more curves on the fault surface (see **Figure 6**). For every value of time, eqn [57] defines a curve on the fault that we call an isochrone.

The integral over the surface in [56] can now be reduced to an integral over the isochrone using standard properties of the delta function

$$\Omega(t,\theta,\phi) = \mu D_i(t)*\int_{l(t)}\frac{\mathrm{d}t}{\mathrm{d}n}\mathrm{d}l \qquad [58]$$

$$= \mu D_i\pi(t)*\int_{l(t)}\frac{v_r}{(1-v_r/c\cos\psi)}\mathrm{d}l \qquad [59]$$

where $l(t)$ is the isochrone and $\mathrm{d}t/\mathrm{d}n=\mathbf{n}\cdot\nabla_{x_0}t=v_r/(1-v_r/c\cos\psi)$ is the derivative of t in the direction perpendicular to the

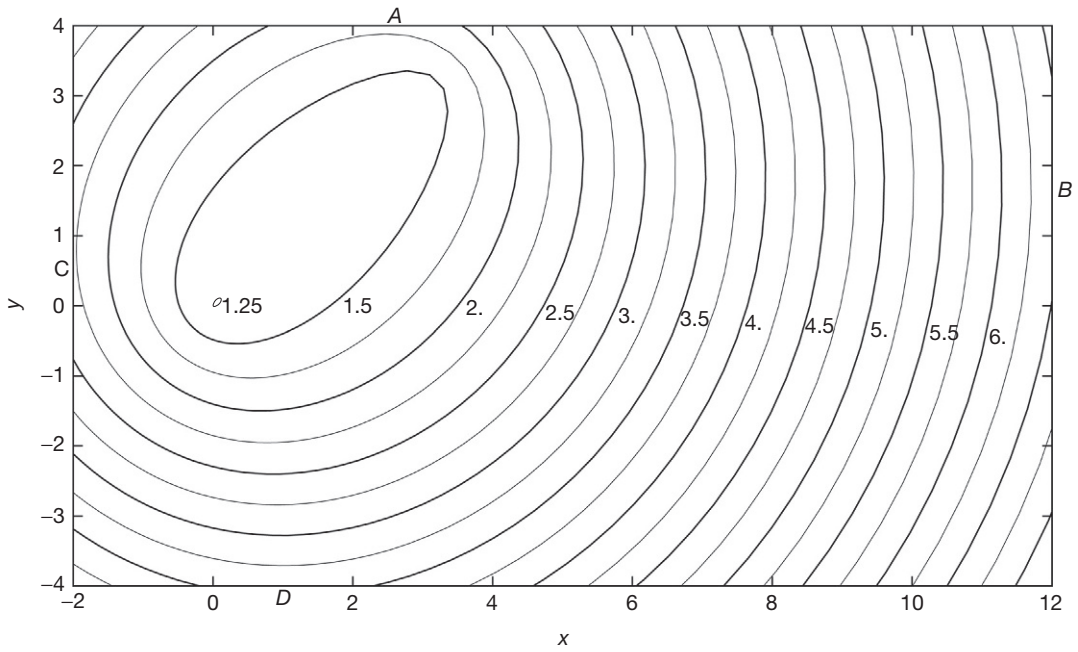


Figure 6 Example of an isochrone. The isochrone was computed for an observer situated at a point of coordinates (3.,3.,1.) in a coordinate system with origin at the rupture initiation point (0.,0.). The vertical axis is out of the fault plane. Rupture starts at $t=0$ at the origin and propagates outward at a speed of 90% the shear wave speed that is 3.5 km s^{-1} in this computation. The signal from the origin arrives at $t=1.25$ s at the observation point. Points A–D denote the location on the border of the fault where isochrones break, producing strong stopping phases.

isochrone. Actually, as shown by Bernard and Madariaga (1984), dt/dn is the local directivity of the radiation from the isochrone. In general, both the isochrone and the normal derivative dt/dn have to be evaluated numerically. The meaning of [58] is simple, the source time function at any instant of time is an integral of the directivity over the isochrone.

The isochrone summation method has been presented in the simplest case here, using the far-field approximation. The method can be used to compute synthetics in the near field too; in that case, changes in the radiation pattern and distance from the source and observer may be included in the computation of the integral [58] without any trouble. The results are excellent as shown by Bernard and Madariaga (1984) who computed synthetic seismograms for a buried circular fault in a half-space and compared them to full numerical synthetics computed by Bouchon (1982). With improvements in computer speed, the use of isochrones to compute synthetics is less attractive, and although the method can be extended to complex media within the ray approximation, most modern computations of synthetics require the appropriate modeling of multipathing, channeled waves, etc. that are difficult to integrate into the isochrone method. Isochrones are still very useful to understand many features of the radiated field and its connection to the rupture process (see, e.g., Bernard and Baumont, 2005).

4.02.5 Crack Models of Seismic Sources

As we have mentioned several times, dislocation models capture some of the most basic geometric properties of seismic sources but have several unphysical features that require careful consideration. For small earthquakes, the kinematic models are generally sufficient, while for larger events – especially in the near field – dislocation models are inadequate because they may not be used to predict high-frequency radiation. A better model of seismic rupture is of course a crack model like Kostrov's self-similar crack. In crack models, slip and stresses are related in a very precise way, so that a finite amount of energy is stored in the vicinity of the crack. Griffith (1920) introduced crack theory using the only requirement that the appearance of a crack in a body does two things: (1) it relaxes stresses and (2) it releases a finite amount of energy. This simple requirement is enough to define many of the properties of cracks, in particular energy balance (see, e.g., Freund, 1989; Kostrov and Das, 1988; Rice, 1980; see also **Chapter 4.03**, Fracture and Frictional Mechanics).

Let us consider the main features of a crack model. Referring to **Figure 7**, we consider a planar fault lying on the plane x, y with normal z. Although the rupture front may have any shape, it is simpler to consider a linear rupture front perpendicular to the x-axis and moving at speed v_r in the positive x direction. Three modes of fracture can be defined with respect to the configuration of **Figure 7**:

- Antiplane, mode III or SH, when slip is in the y direction and stress drops also in this direction, that is, stress σ_{zy} is relaxed by slip.
- Plane, or mode II, when slip is in the x direction and stress drops also in this direction, that is, stress σ_{zx} is relaxed by this mode.
- Opening, or mode I, when the fault opens with a displacement discontinuity in the z direction. In this case, stress σ_{zz} drops to zero.

In natural earthquakes, the opening mode is unlikely to occur at large scales, although it is perfectly possible for very

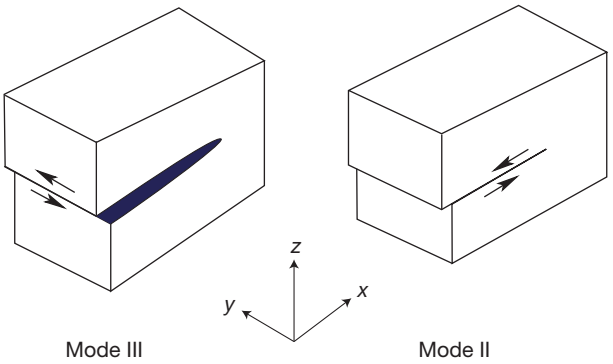


Figure 7 Modes of rupture for shear faulting. Mode III or antiplane mode and mode II or inplane mode may occur at different places on fault boundaries. For general faulting models, both modes occur simultaneously.

small cracks to appear due to stress concentrations, geometric discontinuities of the fault, etc.

For real ruptures, when the rupture front is a curve (or several disjoint ruptures if the source is complex), modes II and III will occur simultaneously on the periphery of the crack. This occurs even in the simple self-similar circular crack model we studied earlier. Fortunately, in homogeneous media, except near-sharp corners and strong discontinuities, the two modes are locally uncoupled, so that most features determined in 2-D carry over to 3-D cracks with little change.

In order to study a two-dimensional crack model, we solve the elastodynamic wave equation together with the following boundary conditions on the $z=0$ plane. For antiplane cracks, mode III:

$$\begin{aligned} \sigma_{zy}(x,0) &= \Delta\sigma \quad \text{for} \quad x < l(t) \\ u_y(x,0) &= 0 \quad \text{for} \quad x > l(t) \end{aligned} \qquad [60]$$

For plane cracks, mode II:

$$\begin{aligned} \sigma_{zx}(x,0) &= \Delta\sigma \quad \text{for} \quad x < l(t) \\ u_x(x,0) &= 0 \quad \text{for} \quad x > l(t) \end{aligned} \qquad [61]$$

where $l(t)=v_r t$ is the current position of the rupture front on the x-axis. These boundary conditions define a mixed boundary value problem that can be solved using complex variable techniques. The solution for arbitrary time variation of $l(t)$ was found for mode III by Kostrov (1966). For plane ruptures, the solution for arbitrary $l(t)$ was found by Freund (1972) (see also Kostrov and Das, 1988). Eshelby (1969) showed that the crack problems have a number of universal features, which are independent of the history of crack propagation and depend only on the instantaneous rupture speed.

4.02.5.1 Rupture Front Mechanics

Since stresses and velocities around a rupture front have universal properties, we can determine them by studying the simpler crack that propagates indefinitely at constant speed. This can be done using a Lorentz transformation of the static elasticity. We are not going to enter in the details of the determination of the solution of the wave equation in moving coordinates; very succinctly, the stress and velocity fields

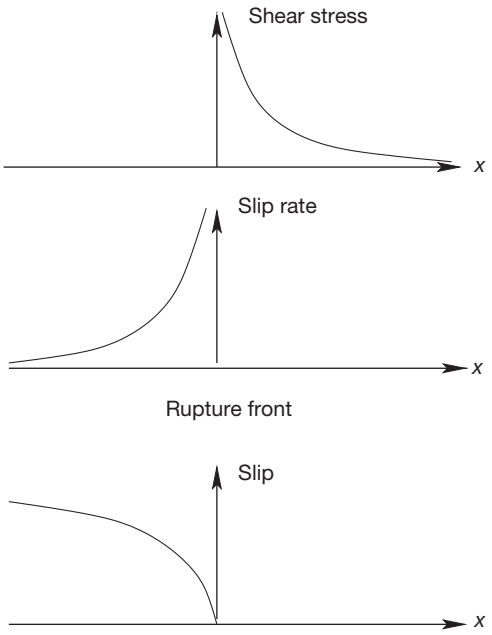


Figure 8 State of stress and slip velocity near the tip of a fracture propagating with a rupture speed less than that of Rayleigh waves.

around the crack tip are related by a nonlinear eigenvalue problem solved by Muskhelishvili in his classical work on complex potentials. There are an infinite number of solutions of the problem, but only one of them ensures a finite energy flow into the crack tip. All other produce no flow or infinite flow.

Along the fault, stress and particle velocities have the universal forms (see Figure 8)

$$\begin{aligned} \sigma(x) &= \frac{K}{\sqrt{2\pi}} \frac{1}{[x-l(t)]^{1/2}} \quad \text{for} \quad x > l(t) \\ \Delta\dot{u}(x) &= \frac{V}{\sqrt{2\pi}} \frac{1}{[x-l(t)]^{1/2}} \quad \text{for} \quad x < l(t) \end{aligned} \qquad [62]$$

and the relations [60] or [61] on the rest of the line. Here, σ stands for either σ_{yz} or σ_{xz} and $\Delta\dot{u}$ for the corresponding slip velocity component in either antiplane or plane fracture modes. In [62], K is the stress concentration, a quantity with units of Pa $m^{1/2}$ that represents the strength of the stress field near the rupture front. V is the dynamic slip rate intensity, which is related to K by

$$K = \frac{\mu}{2v_r}\sqrt{1-v_r^2/\beta^2}\,V \qquad [63]$$

for antiplane cracks, where μ and β are the elastic rigidity and shear wave speed, respectively. For plane cracks, the relation is more complicated

$$K = \frac{\mu}{2v_r}\frac{\beta^2}{v_r^2}\frac{R(v_r)}{\sqrt{1-v_r^2/\beta^2}} \qquad [64]$$

where $R(v_r)$ is the Rayleigh function

$$R(v_r) = 4\sqrt{1-v_r^2/\beta^2}\sqrt{1-v_r^2/\alpha^2} - \left(2-v_r^2/\beta^2\right)^2$$

The complete angular dependence of the stress and particle velocity fields is given by Freund (1989). The inverse squared-root singularities of the form [62] only occur if the rupture velocity is less than the classical limiting rupture speeds: the shear wave velocity β for antiplane ruptures and the Rayleigh wave speed $c_R \approx 0.91\beta$ for inplane cracks. At the terminal speed, the coefficient relating K and V reduces to zero, meaning that at the terminal speed the stress concentration K disappears. A crack running at the terminal speed releases no energy; this speed is thus only possible if there are no rupture resistance and no friction on the fault.

Both experimental evidence and observational evidence cited by Rosakis et al (1999) showed that it is possible for mode II shear cracks to propagate at speeds faster than the shear wave speed in a so-called supershear mode first put in evidence numerically by Andrews (1976). For supershear speeds, the stress and velocity concentration have different dependence with distance than the squared root for subshear faults except for a very particular rupture speed $\sqrt{2}\beta$, sometimes called the Eshelby speed.

The stress field in [62] is infinite at the rupture front; this is a consequence of the idealization that was made to obtain these results: It was assumed that the material around the fault remained elastic even in the immediate vicinity of the rupture front. If more realistic frictional boundary conditions are used instead of the abrupt discontinuities implied by [60] and [61], then the singularity disappears inside a small zone called the rupture process, or slip-weakening zone. Many global features of the crack model can be derived with the simple elastic model studied here; more detailed studies involving a finite slip-weakening zone are only needed to study crack growth in detail.

4.02.5.2 Stress and Velocity Intensity

K and V in [62] have a simpler more fundamental structure that was discovered by Eshelby (1969) for the antiplane case and later extended to plane shear cracks by many authors as reviewed by Freund (1989). Both the dynamic stress intensity factor and the velocity intensity factors can be written in the form

$$\begin{aligned} K &= k(v_r)K^* \\ V &= b(v_r)V^* \end{aligned} \qquad [65]$$

where, for an antiplane crack,

$$k_{III}(v_r) = \sqrt{1 - v_r/\beta}$$

and, for a mode II crack,

$$k_{II}(v_r) \approx 1 - v_r/c_r$$

which is a very good approximation; the exact expression for k_{II} can be found in the books by Freund (1989) and Kostrov and Das (1988).

The factors K^* and V^* in [65] depend only on the load applied to the fault. In the case of an earthquake, they are determined by the stress drop inside those segments of the fault that have already slipped. Since $k(v_r) \to 0$ as the rupture velocity tends to zero, K^* is simply the stress intensity that would prevail if the rupture velocity dropped instantaneously to zero. Thus, K^* is often called the zero-speed stress intensity factor; it depends only on the history of rupture and stress drop inside the broken parts of the fault. Some authors interpret this property of faults mechanics as meaning that rupture fronts have no inertia and their rupture speed can change instantaneously if rupture resistance increases or if the rupture front encounters some geometric discontinuity like a fault jog or a kink.

4.02.5.3 Energy Flow into the Rupture Front

We already mentioned that the stress and slip singularities are a consequence of the requirement that there is a finite energy flow into the rupture front. This energy is used to create new fault surface and is spent in overcoming frictional resistance of the fault. The energy flow into the crack tip was first computed by Kostrov and Nikitin (1970) who provided a very complete discussion of the problem.

The energy flow per unit area of crack advance is usually called G_c

$$G_c = \frac{1}{4v_r} KV \qquad [66]$$

for all the modes, so that for mode III,

$$G_{III} = \frac{1}{2\mu} \frac{1}{\sqrt{1 - v_r^2/\beta^2}} K^2$$

and, for mode II,

$$G_{II} = \frac{1}{2\mu} \frac{v_r^2}{\beta^2} \frac{\sqrt{1 - v_r^2/\beta^2}}{R(v_r)} K^2$$

Let us note that K^2 tends to zero as the rupture velocity v_r approaches the Rayleigh wave speed, so that G_{II} vanishes at the terminal speed.

The crack models are mostly concerned with the local conditions near the edge of the fault as it propagates inside the elastic medium. This is the principal subject of fracture mechanics. In seismology, we are interested not only on the growth of ruptures but also on the generation of seismic waves. Earthquakes are three-dimensional and the finiteness of the source plays a fundamental role in the generation of seismic waves.

4.02.5.4 The Circular Crack

The simplest fault model that can be imagined is a simple circular crack that grows from a point at a constant or variable rupture speed and then stops on the rim of the fault arrested by the presence of unbreakable barriers. The first such simple model was proposed by Madariaga (1976). Although this model is unlikely to represent any actual earthquake, it does quite a good job in explaining many features that are an intrinsic part of seismic sources, most notably the scaling of different measurable quantities, like slip, slip rate, stress change, and energy release. The circular crack problem is posed as a crack problem, which is in terms of stresses not of slip.

We start by a quick study of a simple circular crack from which we derive some of the most fundamental properties of dynamic source models. Let us consider a static circular ('penny-shaped') crack of radius a lying on the x, y plane. Assuming that the fault is loaded by an initial shear stress σ^0_{xz} and that the stress drop

$$\Delta\sigma = \sigma^0_{xz} - \sigma^f_{xz}$$

is uniform, where σ^f_{xz} is the final, residual stress in the fault zone (**Chapters 4.08** and **4.06**), the slip on the fault is given by

$$\Delta u_x(r) = D(r) = \frac{24}{7\pi}\frac{\Delta\sigma}{\mu}\sqrt{a^2 - r^2} \qquad [67]$$

where r is the radial distance from the center of the crack on the (x, y) plane, a is the radius of the crack, and μ is the elastic rigidity of the medium surrounding the crack (see **Figure 9**). Slip in this model has the typical elliptical shape that we associate with cracks and is very different from the constant slip inside the fault assumed in Haskell's model. The taper of the slip near the edges of the crack is of course in agreement with what we discussed about the properties of the elastic fields near the edge of the fault. From [67], we can determine the scalar seismic moment for this circular fault

$$M_0 = \frac{16}{7}\Delta\sigma a^3 \qquad [68]$$

so that the moment is the product of the stress drop times the cube of the fault size. This simple relation is the basis of the seismic scaling law proposed by Aki (1967). The circular crack model has been used to quantify numerous small earthquakes for which the moment was estimated from the amplitude of seismic waves and the source radius was estimated from corner frequencies, aftershock distribution, etc.; the result is that for shallow earthquakes in crustal seismogenic zones like the San Andreas Fault or the North Anatolian Fault in Turkey, stress drops are of the order of 1–10 MPa. For deeper events in subduction zones, stress drops can reach several tens of MPa. Thus, in earthquakes, stresses do not change much, at most a couple of order of magnitudes, while source radius varies over several orders of magnitudes from meters to 100 km or more. It is only in this sense that the usual assertion 'stress drop in earthquakes is constant' should be taken; it actually changes but much less than the other parameters in the scaling law.

Finally, let us take a brief view of the stress field in the vicinity of the fault radius. As expected for crack models, the stress presents stress concentrations of the type [62], that is,

$$\sigma_{xz}(r, \phi) = \frac{(K_{II}\cos\phi + K_{III}\sin\phi)}{\sqrt{r - a}}$$

where (r, ϕ) are polar coordinates on the plane of the circular fault with ϕ being measured from the x-axis. The stress intensity factors are

$$K_{II} = \frac{16}{7\sqrt{\pi}}\frac{1}{1-\nu}\Delta\sigma\sqrt{a} \quad \text{and} \quad K_{III} = \frac{16}{7\sqrt{\pi}}\Delta\sigma\sqrt{a}$$

where ν is Poisson's ratio. It is interesting to note that even if the slip distribution [67] was radially symmetrical, the stress distribution is not. Stress concentration in the mode II direction is stronger than in the antiplane one. As a consequence, if rupture resistance is the same in plane and antiplane modes, a circular crack has an unstable shape. This is clearly observed in fully dynamic simulations where the faults become invariably elongated in the inplane direction.

The creation of a static circular crack with slip [67] inside an elastic medium of rigidity μ produces an elastic energy release

$$\Delta U = \frac{8}{7}\frac{\Delta\sigma^2}{\mu}a^3 \qquad [69]$$

This is the maximum energy that is available for radiation from the dynamic rupture of the circular crack. In reality, the creation of the crack reduces the amount of energy radiated by the energy required to generate the new fault surface.

The simplest physically realistic source model is a circular crack that grows initially at constant speed v_r as predicted by the self-similar Kostrov's model [51] for times less than a/v_r and then evolves spontaneously while it heals. Such a model is usually called quasidynamic because rupture speed is imposed but slip is computed from a properly posed mechanical problem (Madariaga, 1976). There are no simple analytic solutions equivalent to that of Sato and Hirasawa (1973) for quasidynamic cracks. We are forced to use numerical solutions that are actually very simple to obtain using either finite difference or boundary integral equation techniques. The full solution to the circular crack problem is shown in **Figure 10**. Initially, until the sudden arrest of rupture at the final radius a, the slip distribution can be accurately computed using Kostrov's self-similar solution [51]. The stopping of rupture generates strong healing waves that propagate inward from the rim of the fault. These waves are of three types: P-, S-, and Rayleigh waves. Soon after the passage of the Rayleigh waves, slip rate inside the fault decreases to zero and the fault heals. After healing, we assume that frictional forces are sufficiently strong that no slip will

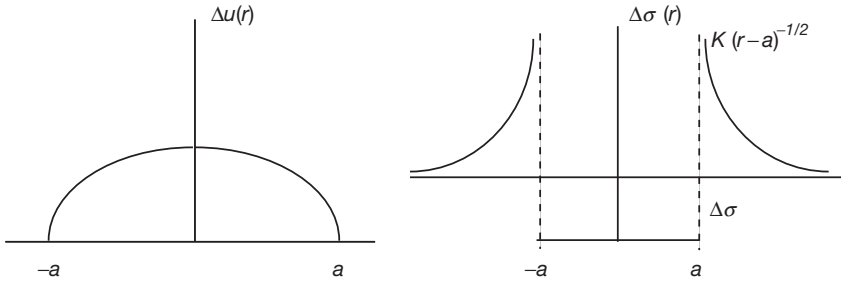


Figure 9 Static circular crack model. On the left, the slip distribution as a function of radius of the fault. On the right, the corresponding shear stress change. Inside the fault, stress drops by a constant amount, and outside the fault, inverse square root stress concentrations appear.

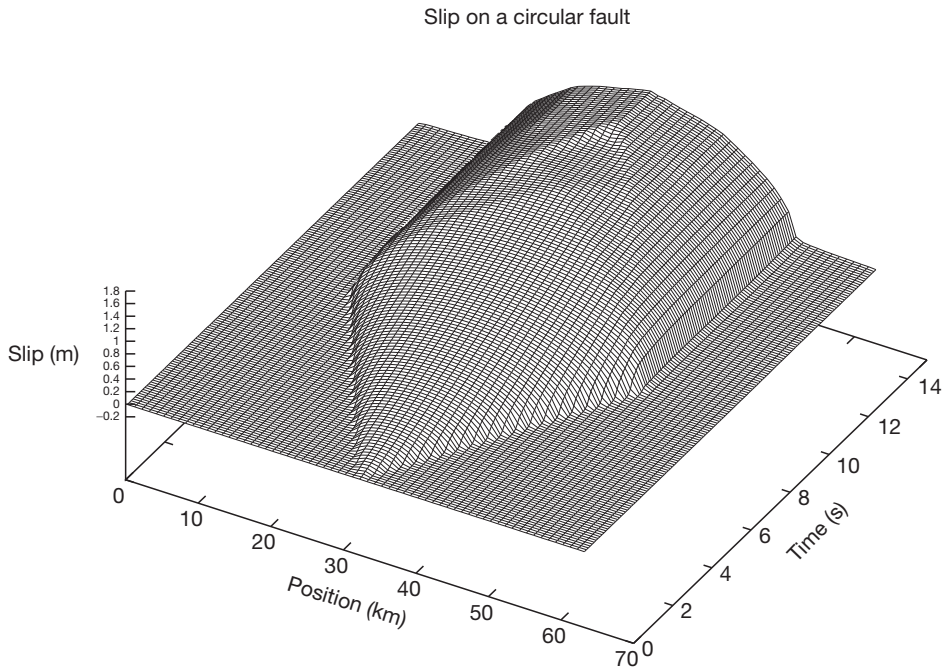


Figure 10 Slip distribution as a function of time and position for a quasidynamic circular crack model. The healing phases, emitted by the border of the fault, are clearly observed. These are the P-, S-, and Rayleigh waves emitted by the circular border of the fault when rupture suddenly arrests.

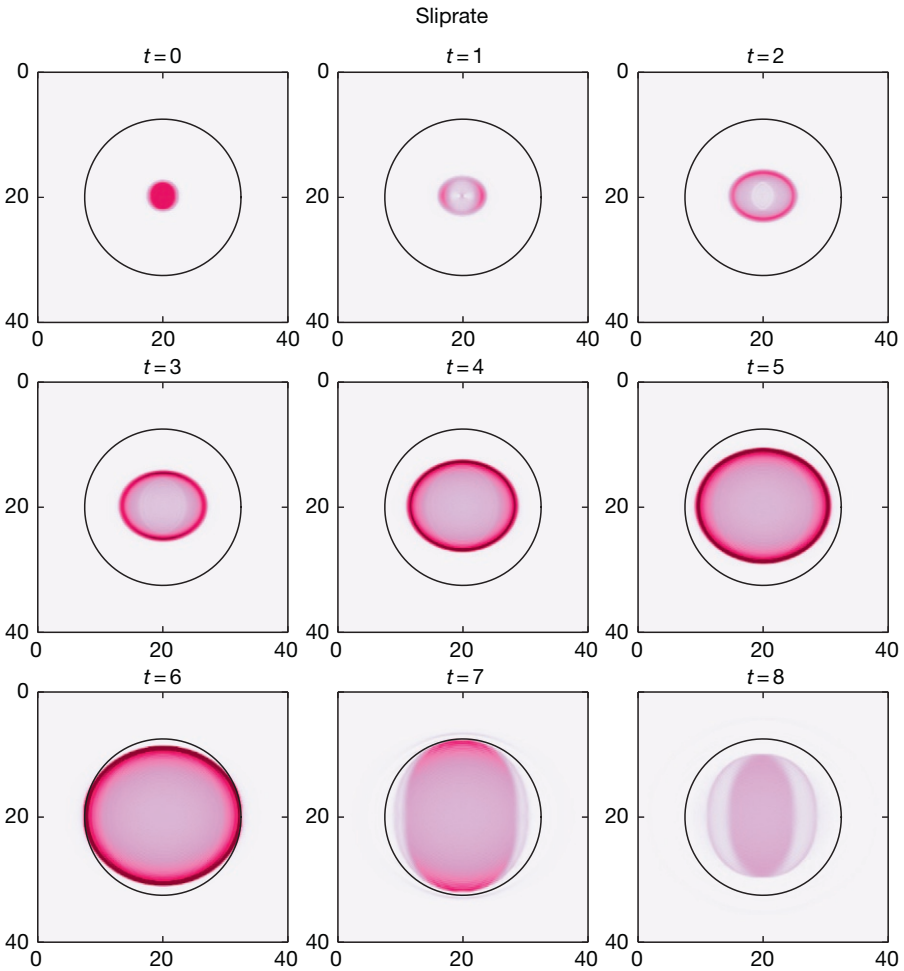


Figure 11 Slip rate on a dynamic circular shear fault. Rupture propagates spontaneously until it reaches an unbreakable circular barrier shown by the black circle. The snapshots depict the slip rate inside the fault at successive instants of time. The darker colors indicate larger values of slip rate, with a scaling ranging from 0 to 1.27 m s^{-1}.

occur until the fault is reloaded. As observed in Figure 10, it is clear that slip and rise time are functions of position on the fault, the rise time being much longer near the center where slip is also larger than near the edges of the fault where slip is clamped. Finally, let us note that the slip after healing is very similar to that of a static circular crack, except that there is an overshoot of slip with respect to the static solution [67]. The overshoot is of course a function of the rupture speed, but its maximum value is of the order of 15% for a rupture speed of 0.75β. The quasidynamic circular crack model was recently studied with a high-resolution numerical method by Kaneko and Shearer (2014). Their results are similar to those of Madariaga (1976) except that the numerical coefficient relating the corner frequency to the radius of the fault has increased. We refer to them for a detailed study of the radiation from the circular crack. Here, we turn to fully dynamic sources.

4.02.5.5 The Dynamic Circular Fault in a Homogenous Medium

The circular crack model was originally proposed by Brune (1970) to explain the omega-squared model of the radiated spectrum. This model is widely used to obtain information about seismic sources, to compute stress change and radiated energy. In this section, we study the spontaneous propagation of rupture away from a small initial asperity that is arrested by an unbreakable circular barrier. This is the simplest realistic source model that we can study. Our study will show the role of rupture propagation, stopping phases, and healing phases for a very simple geometry.

We modeled a circular fault embedded inside a homogeneous elastic medium where P-wave velocity was 6400 m s^{-1}, the P-to-S velocity ratio was 1.73, and the density was 2700 Kg m^{-3}. Stress drop, the difference between initial stress in the fault and the residual stress after slip rate has decreased to zero, was 4.5 MPa. In the dynamic model, rupture propagation is controlled by a slip-weakening friction law (Ida, 1972) with a peak stress of 8 MPa, a slip-weakening distance of 0.2 m, and an energy release rate of $G_c = 0.8$ J m^{-2} (Chapter 4.06).

We modeled the circular fault using a staggered grid finite difference code proposed by Madariaga et al (1998). The space and time steps were 200 m and 0.01 s, respectively. Rupture starts from a small circular asperity where stress is at the static friction level and then propagates spontaneously until it is arrested by an unbreakable barrier of radius of 20 km concentric with the initial asperity. Since we are interested in comparing spontaneous rupture simulations with the quasidynamic circular crack model, we used a stress drop of 4.5 MPa so that the rupture does not become supershear in our simulation. Snapshots of slip rate are shown in Figure 11 at several successive instants of time measured in seconds. We observe that after 2 s, rupture has become elongated in the horizontal direction, which is also the direction of the initial stress. In this direction, mode II prevails. In the transverse direction (vertical), slip is in mode III. Thus, as already remarked by Das (1980) and Day (1982), rupture tends to grow faster in the inplane direction, which is dominated by mode II. At time $t = 6$ s, rupture has reached the unbreakable border of the fault in the inplane direction, and at time $t = 7$ s, the stopping phases generated by the top and bottom edges of the fault have started to move toward the center of the fault. The snapshot at $t = 8$ s shows stopping phases propagating inward from all directions. The slipping patch in darker color is now elongated in the antiplane direction, which is due to slower healing. The duration of the rupture process is slightly longer than 8 s.

Let us now study the radiation from the dynamic circular fault. These circular crack models are widely used to study seismic radiation. How much does *spontaneous rupture* affect the radiation of circular crack models? As expected for a finite rupture model, there are significant variations in the seismic signals radiated in different direction of space. To illustrate this variability, we show in Figure 12 the radiation in two directions: along the mode II direction and along the transverse or mode III direction. The angle θ is the polar angle measured from the normal to the fault plane, and ϕ is the azimuth measured from the direction of slip on the fault. The displacement signals are shown in Figure 12 in two different directions ($\phi = 0°$ and $\phi = 90°$). Both have roughly the same duration of about 10 s but are quite different because one of them ($\phi = 90°$) has larger displacement at the beginning of the signal, while the other is stronger near the end. The reason is that stopping phases, the abrupt change in slope in the figures,

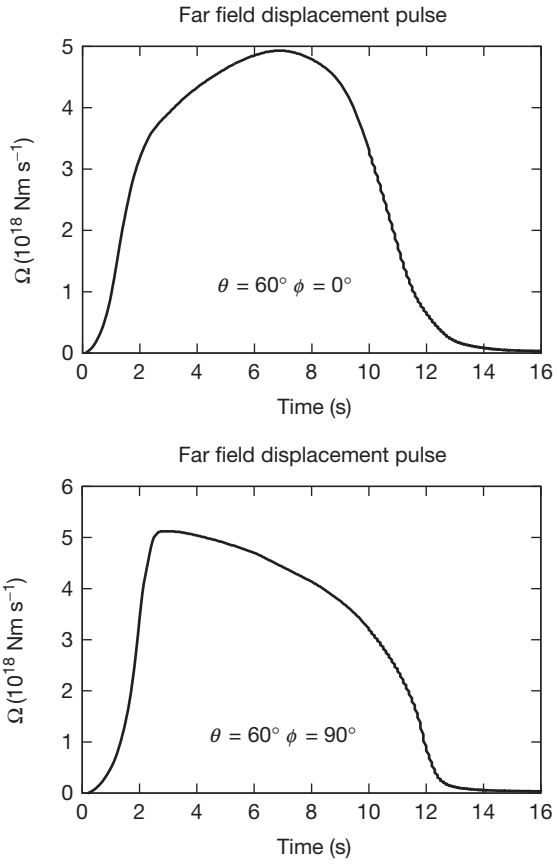


Figure 12 Far-field pulses radiated by the circular crack in two directions of space: θ is the polar angle measured from the axis of the circular crack and ϕ is the azimuth measured from the direction of slip. Thus, $\phi = 0$ is the direction of inplane motion (mode II) and $\phi = 90°$ is the direction of mode III.

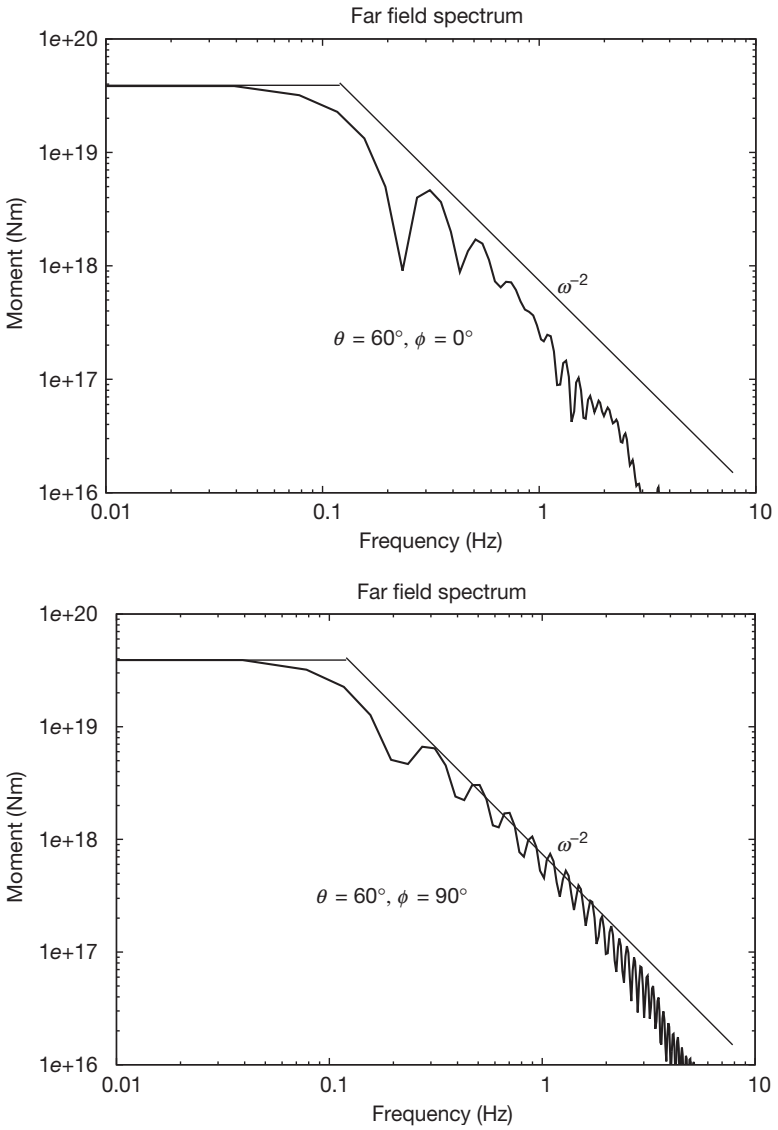


Figure 13 Far-field Fourier spectrum for the two far-field signals shown in **Figure 12**. The spectra have the typical inverse ω-squared decay at high frequencies.

have different relative weights. This difference is in turn due to directivity that is very large in the forward direction ($\phi = 0°$) and is minimal in the transverse direction ($\phi = 90°$). The corresponding amplitude spectra are shown in **Figure 13**. As predicted by Aki (1967) and Brune (1970), these have the typical ω^{-2} spectral decay at high frequencies. The corner frequencies vary from 0.1 to 0.12 Hz. It is not surprising that very different time-domain signals produce similar spectra. The reason is that spectra are dominated by the stopping phases, which carry the information about *rupture arrest*. Although arrest is faster in the inplane direction, the difference in arrival time of the stopping phases is not very large so that the main features of the spectrum are preserved. It is possible to study the variation of corner frequency as a function of the polar angle and azimuth, but the effect of moving the initiation point with respect to the circular border is much more important.

4.02.6 Conclusions

The study of seismic radiation from realistic source models has reached now its maturity. Seismologists have been able to invert the rupture process of a number of earthquakes, and many of the features predicted by simple dynamic source models have been quantified and observed. Foremost among these is the shape of the far-field spectrum, the basic scaling laws relating particle velocity and acceleration to properties of the fault, like size, stress drop, and rupture velocity. The frontier today is the accurate estimation and interpretation of seismic energy and, therefore, the quantification of radiation in terms of the energy balance of seismic sources.

Recent inversions of earthquake slip distributions using kinematic source models have found very complex source

distributions that require an extensive reappraisal of classical source models that were mostly based on Kostrov's model of self-similar circular crack. Ruptures in a fault with heterogeneous stress and rupture resistance distributions follow very tortuous paths. While on the average, the rupture propagates at a sub-Rayleigh speed from one end of the fault to another, in detail, the rupture front can wander in all directions following the areas of strong stress concentration and avoiding those with low stress or high rupture resistance. If this view of earthquake rupture was to be confirmed by future observations (we believe it will be), then many current arguments about earthquake complexity, narrow rupture pulses, and earthquake distributions will be solved, and we may concentrate on the truly interesting problem of determining which features of friction determine that fault stress is always complex under all circumstances.

Acknowledgments

I am greatly indebted to Drs. Shamita Das and John Boatwright for their enlightening comments about the initial version of this manuscript. This research was supported by ANR (Agence nationale de la recherche) under contract S4 ANR-11-BS56-0017.

References

Achenbach JD (1975) *Wave Propagation in Elastic Solids*. Amsterdam: North Holland.

Aki K (1966) Generation and propagation of G waves from the Niigata earthquake of June 16, 1964. Estimation of earthquake movement, released energy and stress-strain drop from G-wave spectrum. *Bulletin of the Earthquake Research Institute* 44: 73–88.

Aki K (1967) Scaling law of seismic spectrum. *Journal of Geophysical Research* 73: 5359–5376.

Aki K and Richards PG (2002) *Quantitative seismology*, 2nd edn. Sausalito, California: University Science Books.

Ampuero J-P and Dahlen FA (2005) Ambiguity of the Moment Tensor. *Bulletin of the Seismological Society of America* 95: 390–400.

Andrews J (1976) Rupture velocity of plane strain shear cracks. *Journal of Geophysical Research* 81: 5679–5687.

Backus G and Mulcahy M (1976) Moment tensors and other phenomenological descriptions of seismic sources. I Continuous displacements. *Geophysical Journal of the Royal Astronomical Society* 46: 321–361.

Ben Menahem A (1961) Radiation of seismic surface waves from finite moving sources. *Bulletin of the Seismological Society of America* 51: 401–435.

Ben Menahem A (1962) Radiation of seismic body waves from finite moving sources. *Journal of Geophysical Research* 67: 345–350.

Ben Menahem A and Singh SJ (1981) *Seismic Waves and Sources*. New York: Springer.

Bernard P and Baumont D (2005) Shear Mach wave characterization for kinematic fault rupture models with constant supershear rupture velocity. *Geophysical Journal International* 162: 431–447.

Bernard P and Madariaga R (1984) A new asymptotic method for the modeling of near-field accelerograms. *Bulletin of the Seismological Society of America* 74: 539–557.

Boatwright J (1980) A spectral theory for circular seismic sources: Simple estimates of source dimension, dynamic stress drop, and radiated seismic energy. *Bulletin of the Seismological Society of America* 70: 1–27.

Bouchon M (1982) The rupture mechanism of the Coyote Lake earthquake of August 6, 1979 inferred from near field data. *Bulletin of the Seismological Society of America* 72: 745–759.

Burridge R and Knopoff L (1964) Body force equivalents for seismic dislocations. *Bulletin of the Seismological Society of America* 54: 1875–1878.

Brune J (1970) Tectonic stress and the spectra of seismic shear waves from earthquakes. *Journal of Geophysical Research* 75: 4997–5009.

Cotton F and Campillo M (1995) Frequency domain inversion of strong motions: Application to the 1992 Landers earthquake. *Journal of Geophysical Research* 100: 3961–3975.

Das S (1980) A numerical method for determination of source time functions for general 3 dimensional rupture propagation. *Geophysical Journal of the Royal Astronomical Society* 62: 591–604.

Dziewonski AM and Woodhouse JH (1983) Studies of the seismic source using normal-mode theory. In: Kanamori H and Boschi E (eds.) *Earthquakes: Observation, Theory, and Interpretation: Notes from the International School of Physics "Enrico Fermi" (1982: Varenna, Italy)*, pp. 45–137. Amsterdam: North-Holland Publication.

Eshelby JD (1956) The continuum theory of lattice defect. In: Seitz F and Turnbull D (eds.) *Progress in Solid State Physics*, vol. 4. New York: Academic Press.

Eshelby JD (1969) The elastic field of a crack extending nonuniformly under general antiplane loading. *Journal of the Mechanics and Physics of Solids* 17: 177–199.

Freund LB (1972) Crack propagation in an elastic solid subjected to general loading. II. Nonuniform rate of extension. *Journal of the Mechanics and Physics of Solids* 20: 141–152.

Freund LB (1989) *Fracture Dynamics*. Cambridge, UK: Cambridge University Press.

Fukuyama E and Madariaga R (1995) Integral equation method for plane crack with arbitrary shape in 3D elastic medium. *Bulletin of the Seismological Society of America* 85: 614–628.

Griffith AA (1920) The phenomenon of rupture and flow in solids. *Philosophical Transactions of the Royal Society (London)* A221: 163–198.

Hanks TC and Kanamori H (1979) Moment magnitude scale. *Journal of Geophysical Research* 84: 2348–2350.

Hartzell SH and Heaton TH (1983) Inversion of strong ground motion and teleseismic waveform data for the fault rupture history of the Imperial Valley, California, earthquake. *Bulletin of the Seismological Society of America* 73: 1553–1583.

Haskell NA (1964) Total energy spectral density of elastic wave radiation from propagating faults. *Bulletin of the Seismological Society of America* 54: 1811–1841.

Haskell NA (1966) Total energy spectral density of elastic wave radiation from propagating faults, Part II. *Bulletin of the Seismological Society of America* 56: 1811–1841.

Hirose S and Achenbach JD (1989) Time-domain boundary element analysis of elastic wave interaction with a crack. *International Journal of Mechanical Engineering* 28: 629–644.

Honda H (1962) Earthquake mechanism and seismic waves. *Journal of Physics of the Earth* 10: 1–98.

Ida Y (1972) Cohesive force across the tip of a longitudinal-shear crack and Griffith's specific surface energy. *Journal of Geophysical Research* 77: 2156–2202.

Kanamori H (1977) The energy release in great earthquakes. *Journal of Geophysical Research* 82: 2921–2987.

Kaneko Y and Shearer P (2014) Seismic source spectra and estimated stress drop derived from cohesive-zone models of circular sub-shear rupture. *Geophysical Journal International*. http://dx.doi.org/10.1093/gji/ggu030 (Online First).

Keylis-Borok VI (1959) On the estimation of the displacement in an earthquake source and of source dimensions. *Annals of Geophysics* 12: 205–214.

Kikuchi M and Kanamori H (1982) Inversion of complex body waves. *Bulletin of the Seismological Society of America* 72: 491–506.

Kikuchi M and Kanamori H (1991) Inversion of complex body waves III. *Bulletin of the Seismological Society of America* 81: 2335–2350.

Knopoff L and Randall MJ (1970) The compensated linear-vector dipole: A possible mechanism for deep earthquakes. *Journal of Geophysical Research* 75: 4957–4963.

Kostrov B (1964) Self-similar problems of propagation of shear cracks. *Journal of Applied Mathematics and Mechanics* 28: 1077–1087.

Kostrov BV (1966) Unsteady propagation of longitudinal shear cracks. *Journal of Applied Mathematics and Mechanics* 30: 1241–1248.

Kostrov B and Das S (1988) *Principles of Earthquake Source Mechanics*. Cambridge: Cambridge University Press.

Kostrov BV and Nikitin L (1970) Some general problems of the mechanics of brittle fracture. *Archiwum Mechaniki Stosowanej* 22: 749–776.

Madariaga R (1976) Dynamics of an expanding circular fault. *Bulletin of the Seismological Society of America* 66: 639–667.

Madariaga R (1978) The dynamic field of Kaskell's rectangular dislocation fault model. *Bulletin of the Seismological Society of America* 68: 869–887.

Madariaga R, Olsen KB, and Archuleta RJ (1998) Modeling dynamic rupture in a 3D earthquake fault model. *Bulletin of the Seismological Society of America* 88: 1182–1197.

Maruyama T (1963) On the force equivalents of dynamical elastic dislocations with reference to the earthquake mechanism. *Bulletin of the Earthquake Research Institute* 41: 467–486.

Nakano H (1923) Notes on the nature of the forces which give rise to the earthquake motions. *Seismological bulletin of the Central Meteorological Observatory of Japan* 1: 92–120.

Rice JR (1980) The mechanics of earthquake rupture. In: Dziewonski AM and Boschi E (eds.) *Physics of the Earth's interior(Proceedings of the International School of Physics 'Enrico Fermi', Course 78, 1979)*, pp. 555–569. Amsterdam: North Holland Publication.

Richards PG (1976) Dynamic motions near an earthquake fault: A three dimensional solution. *Bulletin of the Seismological Society of America* 66: 1–32.

Rosakis AJ, Samudrala O, and Coker D (1999) Cracks faster than the shear wave speed. *Science* 284: 1337–1340.

Sato H and Hirasawa T (1973) Body wave spectra from propagating shear cracks. *Journal of Physics of the Earth* 84: 829–841.

Savage JC (1966) Radiation from a realistic model of faulting. *Bulletin of the Seismological Society of America* 56: 577–592.

Spudich P and Frazer LN (1984) Use of ray theory to calculate high-frequency radiation from earthquake sources having spatially variable rupture velocity and stress drop. *Bulletin of the Seismological Society of America* 74: 2061–2082.

Virieux J (1986) P-SV wave propagation in heterogeneous media: Velocity-stress finite-difference method. *Geophysics* 51: 889–901.

Wald D and Heaton T (1994) Spatial and temporal distribution of slip for the 1992 Landers, California earthquake. *Bulletin of the Seismological Society of America* 84: 668–691.

This page intentionally left blank

4.03 Fracture and Frictional Mechanics: Theory

Y Fialko, University of California San Diego, La Jolla, CA, USA

© 2015 Elsevier B.V. All rights reserved.

4.03.1	**Introduction**	74
4.03.2	**Linear Elastic Fracture Mechanics**	74
4.03.2.1	Singular Crack Models	75
4.03.2.2	Planar Breakdown Zone Models	75
4.03.3	**The Governing Equations**	77
4.03.4	**Exact Solutions for Quasistatic Two-Dimensional Planar Cracks**	77
4.03.5	**Shear Cracks Governed by Rate-and-State Friction Laws**	80
4.03.6	**Dynamic Effects**	81
4.03.7	**Fracture Energy**	83
4.03.8	**Coupling between Elastodynamics and Shear Heating**	85
4.03.9	**Conclusions**	88
Acknowledgments		89
References		89

Nomenclature

ϵ	Small distance	$\Re$	Real part of a complex argument
$\overline{D}$	Average fault slip	R^*	Dynamic length of the process zone at the crack tip
$f_{ij}(\vartheta)$	Azimuthal dependence of a near-tip stress field	r_0	Effective radius of the crack tip
$\hat{T}$	Temperature scale	S	Minor axis of an elliptical cavity
$\overline{w}$	Nondimensional width of the fault slip zone	S_f	Boundary between elastically and inelastically deforming materials
σ_{ij}^0	Stress tensor prior to crack propagation	T	Temperature
σ_{ij}^1	Stress tensor after crack propagation	t	Time
A	Rupture area	U_e	Elastic strain energy
c	Specific heat	U_f	Frictional losses
$D(x)$	Fault slip	U_G	Fracture energy
D_c	Critical weakening displacement	u_i	Displacement vector
D_m	Maximum coseismic displacement	U_l	Potential energy of remotely applied stresses
F	Half-length of a developed part of a crack on which friction has dropped to a residual level σ_d	U_r	Radiated energy
		V_l	Limiting rupture velocity
		V_p	*P*-wave velocity
f_i	Body force	V_r	Rupture velocity
f_s	Coefficient of static friction	V_s	*S*-wave velocity
G	Release rate of mechanical energy (per unit crack length)	w	Half-width of the fault slip zone
		$x, y, z(x_1, x_2, x_3)$	Spatial coordinates
G_c	Effective fracture energy (per unit crack length)	α	Nondimensional factor
		γ	Nondimensional factor
$\Im$	Imaginary part of a complex argument	δ_{ij}	Kronecker delta function
K	Stress intensity factor	ΔS_1	Surface bounding a prospective increment of the process zone at the crack tip in the result of crack growth
K_c	Critical stress intensity factor, fracture toughness		
L	Crack half-length or full length of a self-healing pulse	ΔS_2	Surface bounding a prospective decrement of the process zone at the crack tip in the result of crack growth
L^*	Dynamic pulse length		
L_c	Critical crack length	ΔU_p	Change in the potential energy
M_0	Scalar seismic moment	δW	Work done by external forces
p	Pore fluid pressure	$\Delta\sigma$	Stress drop
Q	Rate of heat generation	ϵ_{ij}	Strain tensor
R	Length of the process zone at the crack tip	θ	Nondimensional temperature
r	Distance to the crack tip	κ	Thermal diffusivity

 http://dx.doi.org/10.1016/B978-0-444-53802-4.00071-3

μ	Shear modulus	σ_n	Normal stress resolved on a fault
ν	Poisson ratio	σ_s	Static strength, or the yield stress in the crack tip process zone
ξ	Dummy variable		
ρ	Density	σ_t	Shear stress resolved on a fault
s	Complex variable	χ	Nondimensional along-crack coordinate
σ_0	Remotely applied stress	ψ	Nondimensional half-length of the developed part of a crack
σ_d	Residual stress on the developed part of a crack	ϕ	Analytic function of a complex argument
σ_{ij}	Stress tensor		

4.03.1 Introduction

Seismic and geodetic observations indicate that most earthquakes are a result of unstable localized shear on quasiplanar faults (Dahlen and Tromp, 1998; Gutenberg and Richter, 1949). Because the thickness of earthquake rupture zones that accommodate slip is much smaller than their characteristic in-plane dimensions, it is natural to idealize earthquake ruptures as shear cracks. Development, propagation, and arrest of shear cracks are subject of the earthquake fracture mechanics. Unlike the engineering fracture mechanics that mainly concerns itself with criteria and details of failure at the tip of tensile cracks propagating through 'intact' solids (Freund, 1998; Lawn, 1993), the earthquake fracture mechanics must consider both the inelastic yielding at the rupture fronts and the evolution of friction (in general, rate- and slip-dependent) on the rest of the slipping surface (see **Chapters 4.04**, **4.05** and **4.06**). Although a distinction is sometimes made between the crack models and friction models of an earthquake source (e.g., Kanamori and Brodsky, 2004), the two processes are intrinsically coupled and should be considered jointly. Note that the shear crack propagation does not necessarily imply creation of a new fault in intact rocks, but also refers to slip on a preexisting (e.g., previously ruptured) interface. Mathematically, the crack growth in unbroken media and on preexisting faults is very similar, provided that the slip surface is planar. While shear cracks in unconfined intact media tend to propagate out of their initial planes, such tendency is suppressed at high confining pressures (e.g., Broberg, 1987; Lockner et al., 1992; Melin, 1986), and in the succeeding text, we limit our attention to planar ruptures.

The redistribution of stress and strain due to a spatially heterogeneous fault slip can be described using either kinematic (displacement-controlled boundary conditions) or dynamic (stress-controlled boundary conditions) approach. Kinematic (e.g., dislocation) models are useful if the fault slip is known or can be inferred with sufficient accuracy, for instance, from seismological, geodetic, or geologic observations (Bilby and Eshelby, 1968; Okada, 1985; Savage, 1998; Steketee, 1958; Vvedenskaya, 1959). Dynamic (fracture mechanics) models potentially have a greater predictive power, as they solve for, rather than stipulate, the details of fault slip for given initial stress conditions (Andrews, 1976a; Ben-Zion and Rice, 1997; Burridge and Halliday, 1971; Day, 1982; Freund, 1979; Madariaga, 1976). The time-dependent boundary conditions on the fault are usually deduced by using constitutive laws that relate kinetic friction to the magnitude and velocity of fault slip, preseismic stress, temperature, and other state variables pertinent to the evolution of stress on a slipping interface. The kinematic and dynamic approaches give rise to identical results for the same slip distribution. Dislocation models are well understood and are widely employed in inversions of seismic and geodetic data for the rupture geometry and spatiotemporal details of slip (e.g., Delouis et al., 2002; Fialko et al., 2005; Hartzell and Heaton, 1983). Fracture mechanics models are intrinsically more complex and less constrained but nonetheless increasingly used to interpret high-quality near-field seismic data (Aochi and Madariaga, 2003; Oglesby and Archuleta, 2001; Peyrat et al., 2001). This chapter will focus on the fracture mechanics approach, and kinematic models will not be discussed. Consequently, the term 'dynamic' will be used to describe time-dependent aspects of rupture for which inertial effects are important, that is, in a meaning that is opposite to 'static' (rather than 'kinematic') descriptions. There exist a number of excellent texts covering the fundamentals of fracture mechanics, with applications to the earthquake source seismology, including Cherepanov (1979), Rice (1980), Rudnicki (1980), Li (1987), Segall (1991), Scholz (2002), and Ben-Zion (2003), among others. This chapter will give an overview of basic aspects of fracture mechanics and some recent theoretical developments.

4.03.2 Linear Elastic Fracture Mechanics

It is well known that structural flaws and discontinuities such as cracks, voids, and inclusions of dissimilar materials give rise to local stress perturbations that may significantly amplify the average applied stress. A classic example is an elliptical cavity in an elastic plate subject to a uniform extension (Inglis, 1913; Lawn, 1993). Provided that the major and minor axes of the cavity are L and S, respectively, the orientation of the remote tensile stress σ_0 is orthogonal to the major axis, and the cavity walls are stress-free, the component of stress parallel to the remote tension assumes a value of $\sigma_0(1+2L/S)=\sigma_0\left(1+2\sqrt{L/r}\right)$ at the cavity tip, where $r=S^2/L$ is the radius of the cavity tip. For an extreme case of a sharp slit, $S/L\to 0$, the stress at the tip scales as $\sigma_0\sqrt{L}/\sqrt{r}$, i.e., becomes unbounded for however small remote loading σ_0. Full analytic solutions for the stress distribution around sharp cracks (see **Section 4.03.4**; also, Freund, 1998; Lawn, 1993; Rice, 1968a) indicate that the stress field indeed has a characteristic square root singularity

$$\sigma_{ij}|_{r\to 0}\approx\frac{K}{\sqrt{2\pi r}}f_{ij}(\vartheta) \qquad [1]$$

where $K\sim O\left(\Delta\sigma\sqrt{L}\right)$ is the stress intensity factor that depends on the crack geometry and loading configuration, $\Delta\sigma$ being the

difference between the far-field stress and stress resolved on the crack walls (hereafter referred to as the stress drop; for an empty crack under remote tension, $\Delta\sigma = \sigma_0$), r is now the distance to the crack tip measured from the crack exterior, and $f_{ij}(\vartheta)$ is a function characterizing the azimuthal dependence of the near-tip stress field (Lawn, 1993; Rice, 1980). Because the governing equations and mathematical structure of solutions are identical for the tensile (mode I), in-plane shear (mode II), and antiplane shear (mode III) cracks (see **Section 4.03.4**), we will use examples of both tensile and shear cracks to highlight universal features of and important differences between the shear and tensile rupture modes. Such a comparison is instructive because many concepts of fracture mechanics have been developed for tensile failure that is most common in engineering applications and subsequently borrowed for seismologic applications that mostly deal with shear failure. Significant differences in the ambient conditions warrant a careful evaluation of the range of applicabilities of basic assumptions behind the fracture mechanics models (e.g., Rubin, 1993).

4.03.2.1 Singular Crack Models

The square root singularity in the stress field (eqn [1]) is a common feature of all crack models assuming a constant stress drop and elastic deformation of the ambient solid, regardless of the mode of failure, and rupture velocity (as long as the latter remains subsonic, see **Section 4.03.5**). While this stress singularity is admissible on thermodynamic grounds (in particular, it is integrable, so that the elastic strain energy is finite in any closed volume containing the crack tip), it is clearly unrealistic as no material is able to support infinite stresses. Theoretical arguments supported by a large number of observations suggest that the assumption of a perfect brittle behavior (i.e., elastic deformation of the unbroken host up to the onset of fracture) is violated in some zone around the crack tip (Atkinson, 1987; Irwin, 1957; Lawn, 1993) (**Figure 1**).

In this zone (commonly referred to as the process or breakdown zone), inelastic yielding prevents stress increases in excess of a certain peak value σ_s that presumably represents the microscopic strength of a material. The size of the process zone r_0 that corresponds to equilibrium (i.e., a crack on a verge of propagating) may be interpreted as the effective radius of the crack tip. Outside of the process zone (at distances $r > r_0$),

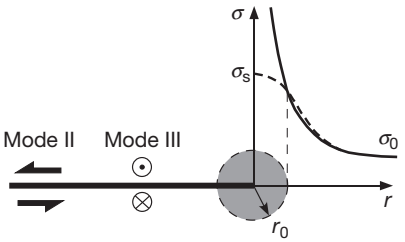


Figure 1 An idealized view of stress variations and yielding at the crack tip. r_0 is the characteristic dimension of a inelastic zone in which stresses exceed the yield threshold σ_s. Solid curve shows the theoretically predicted stress increase with a characteristic $1/\sqrt{r}$ singularity. The theoretical prediction breaks down at distances $r < r_0$ but may be adequate for $r > r_0$.

stresses are below the failure envelope, and the material deforms elastically. If the size of the equilibrium process zone is negligible compared to the crack length, as well as any other characteristic dimensions (e.g., those of an encompassing body), a condition termed small-scale yielding (Rice, 1968a) is achieved, such that the stress and strain fields in the intact material are not appreciably different from those predicted in the absence of a process zone. This is the realm of the linear elastic fracture mechanics (LEFM). According to LEFM, the near-tip ($r_0 < r \ll L$) stress field is completely specified by the scalar multiplier on the singular stress field, the stress intensity factor K (eqn [1]), which can be found by solving an elastic problem for a prescribed crack geometry and loading conditions. The crack propagation occurs when the stress intensity factor exceeds a critical value, $K > K_c$. The critical stress intensity factor, or the fracture toughness K_c, is believed to be a material property, independent of the crack length and loading configuration (although fracture properties may vary for different modes of failure, e.g., K_{Ic} and K_{IIc}, similar to differences between macroscopic tensile and shear strengths). To the extent that the microscopic yield strength σ_s and the effective equilibrium curvature of the crack tip, r_0, may be deemed physical properties, the fracture toughness K_c may be interpreted as a product $\sigma_s\sqrt{r_0} = \text{const}$.

The critical stress intensity factor is a local fracture criterion, as it quantifies the magnitude of the near-tip stress field on the verge of failure. However, it can be readily related to global parameters characterizing changes in the elastic strain energy ∂U_e and potential energy of applied stresses ∂U_l in the entire body due to a crack extension ∂L, such as the energy release rate $G_c = -(\partial U_e + \partial U_l)/\partial L$,

$$G_c = \frac{\alpha K_c^2}{2\mu} \qquad [2]$$

where μ is the shear modulus of an intact material and α assumes values of $(1-\nu)$ and unity for mode II and mode III loading, respectively (Irwin, 1957). For ideally brittle materials, the energy release rate may be in turn associated with the specific surface energy spent on breaking the intermolecular bonds (Griffith, 1920; Lawn, 1993). Further analysis of the breakdown process at the crack tip requires explicit consideration of the details of stress concentration in the tip region.

4.03.2.2 Planar Breakdown Zone Models

A simple yet powerful extension of the LEFM formulation is the displacement-weakening model, which postulates that (i) the breakdown process is confined to the crack plane, (ii) inelastic deformation begins when stresses at the crack tip reach some critical level σ_s, and (iii) yielding is complete when the crack wall displacement exceeds some critical value D_c (Barenblatt, 1959; Dugdale, 1960; Leonov and Panasyuk, 1959). In case of tensile (mode I) cracks, σ_s represents the local tensile strength in the breakdown zone, and D_c is the critical opening displacement beyond which there is no cohesion between the crack walls. In case of shear (modes II and III) cracks (**Figure 2**), σ_s represents either the shear strength (for ruptures propagating through an intact solid) or the peak static friction (for ruptures propagating along a preexisting fault), and D_c is the

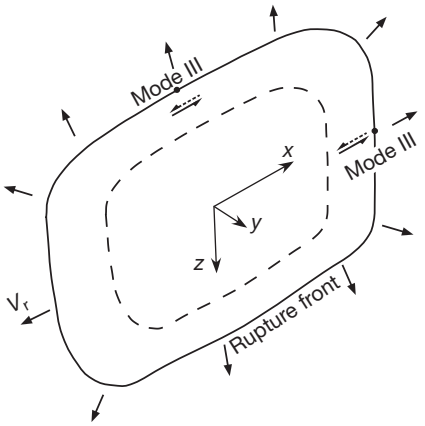


Figure 2 Schematic view of a shear rupture expanding at a constant velocity V. Conditions corresponding to the two-dimensional mode II and mode III loading are approximately satisfied at the rupture fronts orthogonal and parallel to the local slip vector, respectively. For a cracklike rupture, slip occurs on the entire area bounded by the rupture front. For a pulse-like rupture, slip is confined to an area between the rupture front (solid line) and a trailing healing front (dashed line).

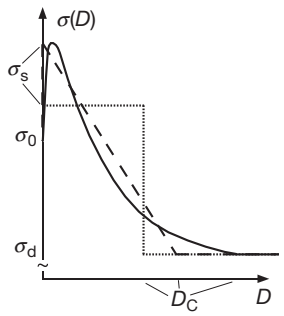


Figure 3 Possible dependence of the breakdown stress σ on the crack wall displacement D within the crack tip process zone. Solid line: a schematic representation of the experimentally measured displacement-weakening relations for rocks (Hashida et al., 1993; Li, 1987). Dashed and dotted lines: approximations of the displacement-weakening relation assuming no dependence of the yield stress on D (dotted line) and linear weakening (dashed line). The area beneath all curves is the same and equals to the fracture energy $G_c = \alpha\sigma_s D_c$, where α is a numerical factor that depends on the particular form of the displacement-weakening relationship. For example, $\alpha = 1$ for the constant yield stress model, and $\alpha = 0.5$ for the linear weakening model.

slip-weakening distance at which a transition to the kinetic friction is complete (Ida, 1972; Palmer and Rice, 1973). Under these assumptions, the fracture energy may be defined as work required to evolve stresses acting on the crack plane from σ_s to the residual value σ_d and is of the order of $(\sigma_s - \sigma_d)D_c$, depending on the details of the displacement-weakening relation (Figure 3). For an ideal brittle fracture that involves severing of intermolecular bonds ahead of the crack tip, σ_s may approach theoretical strength, $\mu/10 \sim O(10^9\text{–}10^{10}\ \text{Pa})$, and D_c may be of the order of the crystal lattice spacing (10^{-10}–10^{-9} m), yielding $G_c \sim O(1\ \text{J m}^{-2})$. This is close to the experimentally measured fracture energies of highly brittle crystals and glasses (Griffith, 1920; Lawn, 1993). At the same time, laboratory measurements of polycrystalline aggregates (such as rocks, ceramics, and metals) reveal much higher fracture energies ranging from $10\text{–}10^2\ \text{J m}^{-2}$ (for tensile failure) to $10^3\ \text{J m}^{-2}$ (for shear failure), presumably reflecting dependence of the effective D_c on the material microstructure (e.g., texture, grain size, and distribution of inhomogeneities) and the breakdown mechanism (e.g., grain unlocking, ligamentary bridging, plasticity, off-plane yielding, or some other deviations from the ideally brittle behavior) (Fialko and Rubin, 1997; Hashida et al., 1993; Li, 1987). Indirect inferences of fracture energies associated with *in situ* propagation of tensile cracks such as man-made hydrofractures and magmatic dikes (Rubin, 1993) and earthquake ruptures (Abercrombie and Rice, 2005; Chester et al., 2005; Husseini, 1977; Wilson et al., 2004) reveal significantly larger fracture energies compared with those measured in the laboratory experiments. This disagreement may be indicative of some scaling of fracture energy with the rupture size or rupture history and perhaps the significance of dynamic effects on cracking and damage that are not captured in laboratory tests. In the context of an in-plane breakdown zone model, the inferred scale dependence of fracture energy is usually interpreted in terms of increases in the effective weakening displacement D_c with the rupture length L. Possible mechanisms of scale dependence of D_c include continued degradation of the dynamic fault strength with slip (e.g., due to thermal or rheological weakening) (Abercrombie and Rice, 2005; Brodsky and Kanamori, 2001; Di Toro et al., 2004; Lachenbruch, 1980) or fractal nature of the fault roughness (e.g., Barton, 1971; Power and Tullis, 1995). It should be pointed out that the common interpretation of the scale dependence of G_c in terms of the scale dependence of D_c may be too simplistic, as the assumption of a thin process zone unlikely holds under *in situ* stress conditions for either tensile or shear cracks, as suggested by both theoretical arguments (Fialko and Rubin, 1997; Poliakov et al., 2002; Rice et al., 2005; Rubin, 1993) and observations (Chester et al., 2005; Fialko, 2004b; Fialko et al., 2002; Manighetti et al., 2001; Wilson et al., 2004). A more detailed quantitative treatment of the energetics of fracture and breakdown processes is presented in **Section 4.03.7**.

Nonsingular crack models postulating a thin in-plane process zone are mathematically appealing because they allow one to treat the fracture problem as an elastic one by removing nonlinearities associated with failure from the governing equations and incorporating them into boundary conditions on a crack plane. For example, elastic solutions can be readily used to determine the size of the process zone at the crack tip. Dimensional arguments indicate that the strain associated with the material breakdown is of the order of D_c/R, where R is the characteristic size of the breakdown zone. This strain must be of the order of the elastic strain associated with the strength drop, $(\sigma_s - \sigma_d)/\mu$, implying

$$R = \gamma \frac{\mu}{\sigma_s - \sigma_d} D_c \qquad [3]$$

where γ is a nondimensional factor of the order of unity that depends on the displacement-weakening relation (e.g., as shown in Figure 3). The small-scale yielding condition requires $R \ll L$. Parallels may be drawn between the size of the breakdown zone R in the displacement-weakening model and the critical radius of the crack tip r_0 in the LEFM model

(Figure 1; also, see Khazan and Fialko, 1995), in that both parameters demarcate the near-tip region that undergoes failure from the rest of the material in which the deformation is essentially elastic. On a larger scale, the assumption of elasticity postulates a direct proportionality between the stress drop $\Delta\sigma$ and the average fault slip $\overline{D}$:

$$\overline{D} \sim \frac{\Delta\sigma}{\mu} L \qquad [4]$$

where L is the characteristic fault dimension (for isometric ruptures) or the least dimension of the slip area (in case of ruptures having irregular shape or high aspect ratios). The assumption of a predominantly elastic behavior of the Earth's crust appears to be in a good agreement with seismic (Aki and Richards, 1980; Gutenberg and Richter, 1949; Vvedenskaya, 1959) and geodetic (Fialko, 2004b; Reid, 1910) observations of the instantaneous response of crustal rocks to major earthquakes. The brittle–elastic model gives rise to several fundamental scaling relationships used in earthquake seismology. For example, the scalar seismic moment M_0 is a measure of the earthquake size:

$$M_0 = \overline{D} A \mu \qquad [5]$$

where A is the rupture area (Aki, 1967; Kostrov, 1974; also, see Chapter 4.02). The seismic moment is related to a coseismic change in the potential energy of elastic deformation ΔU_e (e.g., Kanamori, 2004),

$$\Delta U_e = \frac{\Delta\sigma}{2\mu} M_0 \qquad [6]$$

For isometric ruptures, $A \sim L^2$, and by combining eqns [4] and [5], one obtains a well-known scaling between the scalar seismic moment, the rupture size, and the stress drop (e.g., Dahlen and Tromp, 1998; Kanamori and Anderson, 1975, Chapter 4.02):

$$M_0 \sim \Delta\sigma L^3 \qquad [7]$$

4.03.3 The Governing Equations

Consider a three-dimensional medium which points are uniquely characterized by Cartesian coordinates x_i ($i = 1, 2, 3$) in some reference state prior to the fault-induced deformation. Fault slip gives rise to a displacement field u_i. The displacements u_i are in general continuous (differentiable), except across the slipped part of the fault. The strain tensor is related to the displacement gradients as follows:

$$\epsilon_{ij} = \frac{1}{2}\left(u_{i,j} + u_{j,i}\right) \qquad [8]$$

where the comma operator as usual denotes differentiation with respect to spatial coordinates, $a_{,i} = \partial a/\partial x_i$. Equation [8] assumes that strains are small ($\epsilon_{ij} \ll 1$), so that terms that are quadratic in displacement gradients can be safely neglected (e.g., Landau and Lifshitz, 1986; Malvern, 1969). The assumption of an infinitesimal strain implies no difference between the material (Lagrangian) and spatial (Eulerian) reference frames. Typical strains associated with earthquake ruptures are of the order of 10^{-4}–10^{-5} (Kanamori and Anderson, 1975; Scholz, 2002) so that the infinitesimal strain approximation is likely justified. For sufficiently small strain changes (with respect to some reference state), laboratory data and theoretical arguments suggest a linear dependence of strain perturbation on the causal stress change. For an isotropic homogeneous elastic material, the corresponding relationship between stresses and strains is given by Hooke's law (e.g., Landau and Lifshitz, 1986; Timoshenko and Goodier, 1970):

$$\epsilon_{ij} = \frac{1}{2(1+\nu)\mu}\left[(1+\nu)\sigma_{ij} - \nu\delta_{ij}\sigma_{kk}\right] \qquad [9]$$

where μ and ν are the shear modulus and Poisson's ratio, respectively, δ_{ij} is the Kronecker delta function, and repeating indexes imply summation.

Conservation of a linear momentum in continuous media gives rise to the Navier–Cauchy equations of equilibrium (Malvern, 1969):

$$\sigma_{ij,\,j} + f_i = \frac{\partial^2 u_i}{\partial t^2} \qquad [10]$$

where f_i is the body force (e.g., due to gravity) and t is time. Due to the linearity of equilibrium equations [10], it is possible to represent the full stress tensor as a superposition of some background (e.g., lithostatic and regional tectonic) stress and a perturbation due to fault slip, such that the latter satisfies a homogeneous case of eqns [10]. Unless noted otherwise, in the succeeding text, we assume that the stress tensor σ_{ij} denotes only stress perturbations due to fault displacements. For convenience, the indicial nomenclature for spatial coordinates (x_1, x_2, x_3) will be used interchangeably with the traditional component notation (x, y, z) throughout the rest of the text.

In order to close the problem formulation, some constitutive law relating slip D, slip velocity $\partial D/\partial t$, resolved shear and normal stresses, σ_t and σ_n, pore fluid pressure p, temperature T, etc. needs to be prescribed on surfaces that violate assumptions of continuity, for example, faults and cracks. Examples are the Mohr–Coulomb (Byerlee, 1978) and rate-and-state friction (see Chapter 4.04), flash melting (Molinari et al., 1999; Rice, 2006), thermal pressurization (Andrews, 2002; Segall and Rice, 1995; Sibson, 1973), viscous rheology (Fialko and Khazan, 2005), etc.

4.03.4 Exact Solutions for Quasistatic Two-Dimensional Planar Cracks

In case of two-dimensional deformation (e.g., plane or antiplane strain or plane stress), elastic solutions for stresses and displacements can be generally expressed in terms of two analytic functions of a complex variable $\zeta = x + iy$ (Kolosoff, 1909; Muskhelishvili, 1953). For simplicity, here, we consider loading that is symmetric about the center of the crack $x = 0$. In this case, both stresses σ_{ij} and displacements u_i can be expressed through a single analytic function $\phi(\zeta)$ (Westergaard, 1939). First, we demonstrate that the mathematical structure of solutions for stresses and displacements in the crack plane ($y = 0$) is identical for tensile (mode I), in-plane shear (mode II), and antiplane shear (mode III) cracks. In particular, for tensile

(mode I) cracks, stresses and displacements can be found from $\phi(\zeta)$ as (e.g., Khazan and Fialko, 2001; Muskhelishvili, 1953)

$$\sigma_{\gamma\gamma} = 2(\Re\phi' + \gamma\Im\phi'') \quad [11]$$

$$\sigma_{x\gamma} = -2\gamma\Re\phi'' \quad [12]$$

$$u_\gamma = \frac{1}{\mu}(2(1-\nu)\Im\phi - \gamma\Re\phi') \quad [13]$$

where $\Re$ and $\Im$ denote the real and imaginary parts of a complex argument. For plane strain shear (mode II) cracks, conditions of symmetry imply that $\sigma_{\gamma\gamma}=0$ on the crack plane ($\gamma=0$), and the corresponding equilibrium equations are

$$\sigma_{x\gamma} = -2(\Im\phi' + \gamma\Re\phi'') \quad [14]$$

$$u_x = \frac{1}{\mu}(2(1-\nu)\Re\phi + \Im\phi) \quad [15]$$

Upon making a substitution $\phi = -i\phi_1$, one can see that the unknown shear stress $\sigma_{x\gamma}$ and displacement u_x on the crack plane may be expressed through a new function ϕ_1 in the same manner as the normal stress $\sigma_{\gamma\gamma}$ and displacement u_γ are expressed through ϕ (eqns [11] and [13]),

$$\sigma_{x\gamma} = 2\left(\Re\phi_1' - \gamma\Im\phi_1''\right) \quad [16]$$

$$u_x = \frac{1}{\mu}\left(2(1-\nu)\Im\phi_1 - \gamma\Re\phi_1'\right) \quad [17]$$

For antiplane shear (mode III) cracks, expressions for the relevant stress and displacement components are

$$\sigma_{yz} = -\Im\phi' \quad [18]$$

$$u_z = \frac{1}{\mu}\Re\phi \quad [19]$$

Upon making a substitution $\phi = -2i\phi_2$, one can see that the dependence of the unknown quantities σ_{yz} and $(1-\nu)u_z$ on the analytic function ϕ_2 is analogous to the dependence of $\sigma_{x\gamma}$ and u_x on ϕ_1 obtained for the in-plane shear crack for $\gamma=0$ (eqns [16]–[17]). This analogy mandates that the mathematical structure of solutions for the tensile, in-plane, and antiplane components of stress for the corresponding crack modes is identical, provided that the boundary conditions on the crack plane (the along-crack distribution of the driving stress and the displacement-weakening relationship) are analogous. Solutions for the crack wall displacements are also identical for different modes, although expressions for displacements for mode III cracks will differ from those for mode I and mode II cracks by a factor of $(1-\nu)$. Hence, we focus on a particular case of an in-pane shear (mode II) crack. The boundary conditions for the potential function ϕ_1 are as follows:

$$\Re\phi_1' = \sigma(x)/2 \quad \text{for } |x| < L \quad [20]$$

$$\Im\phi_1' = 0 \quad \text{for } |x| > L \quad [21]$$

$$\phi_1' \to \sigma_0/2 \quad \text{for } |\zeta| \to \infty \quad [22]$$

where $\sigma(x)$ is the distribution of shear stress on the crack surface and σ_0 is the applied shear stress at infinity (hereafter referred to as prestress). The boundary condition [21] postulates no slip beyond the rupture front (see eqn [17]). An explicit solution for the function ϕ_1 that is analytic in the upper half plane and satisfies the boundary conditions [20]–[22] is given by the Keldysh–Sedov formula (Gakhov, 1966; Khazan and Fialko, 1995, 2001; Muskhelishvili, 1953):

$$\phi_1'(\zeta) = \frac{1}{2\pi i}\left(\frac{\zeta+L}{\zeta-L}\right)^{1/2}\int_{-L}^{L} d\xi\left(\frac{\xi-L}{\xi+L}\right)^{1/2}\frac{\sigma(\xi)}{\xi-\zeta} + \frac{\sigma_0}{2}\left(\frac{\zeta+L}{\zeta-L}\right)^{1/2} + \frac{C}{\left(\zeta^2-L^2\right)^{1/2}} \quad [23]$$

where C is an arbitrary constant. From eqn [16], an asymptotic behavior of ϕ_1' at infinity is

$$\lim_{|\zeta|\to\infty}\phi_1'(\zeta) = -\frac{1}{2\pi i|\zeta|}\int_{-L}^{L} d\xi\left(\frac{\xi-L}{\xi+L}\right)^{1/2}\sigma(\xi) + \frac{\sigma_0}{2} + \frac{L\sigma_0 + C}{|\zeta|} \quad [24]$$

One can readily determine the unknown constant C from eqn [16] by satisfying the boundary condition [22]. The final expression for the derivative of analytic function ϕ_1 is

$$\phi_1'(\zeta) = \frac{1}{2\pi i}\left(\frac{\zeta+L}{\zeta-L}\right)^{1/2}\int_{-L}^{L} d\xi\left(\frac{\xi-L}{\xi+L}\right)^{1/2}\frac{\sigma(\xi)}{\xi-\zeta} + \frac{\sigma_0}{2}\left(\frac{\zeta+L}{\zeta-L}\right)^{1/2} + \frac{1}{2\pi\left(\zeta^2-L^2\right)^{1/2}}\int_{-L}^{L} d\xi\frac{\sigma(\xi)-\sigma_0}{\left(L^2-\xi^2\right)^{1/2}} \quad [25]$$

For any physically admissible failure model, the maximum stress within the slipped region is bounded by the yield strength, $\sigma_{x\gamma} < \sigma_s$ for $|x| < L$. Furthermore, it is reasonable to expect that $\sigma_{x\gamma} \to \sigma_s$ as $|x| \to L$. The shear stress $\sigma_{x\gamma}(\zeta)$ in the 'locked' region ahead of the rupture front is then given by the real part of ϕ_1 (see eqn [16]). Sufficiently close to the crack tip (i.e., for $\zeta = L + \epsilon$, such that $\Im\epsilon = 0$, $0 < \epsilon \ll L$), we obtain

$$\sigma_{x\gamma}(L+\epsilon) = \frac{1}{\pi}\left(\frac{L}{2\epsilon}\right)^{1/2}\int_{-L}^{L} d\xi\frac{\sigma(\xi)-\sigma_0}{\left(L^2-\xi^2\right)^{1/2}} + \sigma_s + O\left(\epsilon^{1/2}\right) \quad [26]$$

The first term on the right-hand side of eqn [26] is of the order of $1/\sqrt{\epsilon}$ and represents the LEFM approximation (see Section 4.03.2). By comparing eqns [26] and [1], one can formally introduce the stress intensity factor,

$$K = \frac{1}{\pi}\left(\frac{L}{2\epsilon}\right)^{1/2}\int_{-L}^{L} d\xi\frac{\sigma(\xi)-\sigma_0}{\left(L^2-\xi^2\right)^{1/2}} \quad [27]$$

which exhibits the expected scaling $K \propto \Delta\sigma\sqrt{L}$. Physically plausible crack models require that stresses are finite everywhere. From eqns [26] and [27], the requirement of the absence of a stress singularity is met if (and only if)

$$\int_{-L}^{L} d\xi\frac{\sigma(\xi)-\sigma_0}{\left(L^2-\xi^2\right)^{1/2}} = 0 \quad [28]$$

Any realistic distribution of the driving stress resolved the crack surface must satisfy the integral constraint [28]. This implies that the driving stress $(\sigma(x)-\sigma_0)$ must change sign along the crack ($|x| < L$). For example, the stress drop $(\sigma(x) < \sigma_0)$ in the central part of the crack needs to be balanced by the material strength or the high transient friction $(\sigma_s \geq \sigma(x) > \sigma_0)$ within the process zone.

Displacements of the crack walls $u_x(x)$ corresponding to the instantaneous shear stress $\sigma(x)$ can be found by differentiating equation [17] for $\gamma=0$,

$$\frac{du_x}{dx} = \frac{2(1-\nu)}{\mu}\Im\phi_1' \qquad [29]$$

and making use of expression [26] to integrate the resulting differential equation [29] with the initial condition $u_x(-L)=0$. The respective solution is

$$D(x) = \frac{2(1-\nu)}{\pi\mu}\int_x^L (L^2-\xi^2)^{1/2} d\xi \int_{-L}^{L} \frac{\sigma(t)dt}{(L^2-t^2)^{1/2}(t-\xi)} \qquad [30]$$

where $D(x)=2u_x(x)$ is slip between the crack walls. The problem is closed by specifying a constitutive relationship between slip D and kinetic friction σ. In general, such a relationship may include dependence of friction on slip rate, slip history, local temperature, and other state variables (Dieterich, 1979; Blanpied et al., 1995; Ruina, 1983, **Chapters 4.04, 4.05** and **4.06**). To gain a further analytic insight, here, we consider a simple slip-weakening relationship

$$\begin{aligned} \sigma(x) &= \sigma_s \quad \text{for} D(x) < D_c \\ \sigma(x) &= \sigma_d \quad \text{for} D(x) > D_c \end{aligned} \qquad [31]$$

where σ_s is the yield strength or static friction in the process zone, σ_d is the residual kinetic friction on the developed part of the crack, and D_c is the critical slip-weakening displacement corresponding to a transition from σ_s to σ_d (see **Figure 3**). The size of the process zone, R, is defined by a requirement that the fault slip is subcritical within the process zone, $D(|F|)=D_c$, where $F=L-R$ is the length of the developed part of a crack on which the friction has dropped to a residual value σ_d. **Figure 4** illustrates the geometry of the problem. In case of a piece-wise constant distribution of shear stress along the crack $\sigma(x)$ given by eqn [32], as shown in **Figure 4**, expressions [28] and [30] can be readily integrated to provide closed form analytic solutions. In particular, integration of expression [28] allows one to determine a relative size of the process zone with respect to the crack half-length:

$$\frac{R}{L} = 2\sin^2\left(\frac{\pi}{4}\frac{\sigma_0-\sigma_d}{\sigma_s-\sigma_d}\right) \qquad [32]$$

Evaluation of integral [30] gives rise to

$$D(x) = \frac{2(1-\nu)L}{\pi\mu}(\sigma_0-\sigma_d)I(\chi,\psi) \qquad [33]$$

where $\chi=x/L$ and $\psi=F/L$ are the nondimensional along-crack coordinate and the half-length of the developed part of the crack, respectively, and function I is given by the following equation:

$$\begin{aligned} I(U,V) = (V+U)\log\left|\frac{1+UV+\sqrt{(1-U^2)(1-V^2)}}{V+U}\right| \\ +(V-U)\log\left|\frac{1-UV+\sqrt{(1-U^2)(1-V^2)}}{V-U}\right| \end{aligned} \qquad [34]$$

At the base of the process zone, $\chi=\psi$, eqn [33] reduces to

$$D(F) = \frac{4(1-\nu)}{\pi\mu}F(\sigma_s-\sigma_d)\log\frac{1}{\psi} \qquad [35]$$

For a crack that is on the verge of propagating, slip at the base of the process zone ($x=F$) equals the critical weakening displacement D_c. A simple dimensional analysis of eqn [35] then reveals a characteristic length scale L_c:

$$L_c = \frac{\pi D_c}{4(1-\nu)}\frac{\mu}{\sigma_s-\sigma_d} \qquad [36]$$

Using eqn [36], eqn [35] may be written

$$\frac{L_c}{F} = \log\frac{L}{F} \qquad [37]$$

A further rearrangement of eqn [37] yields an explicit expression for the length of the process zone R:

$$R = F\left(\exp\left(\frac{L_c}{F}\right)-1\right) \qquad [38]$$

It is instructive to consider two end-member cases, the long crack limit, $F \gg L_c$, and the microcrack limit, $F \ll L_c$. In case of sufficiently long cracks ($F \gg L_c$), an asymptotic expansion of eqn [38] (i.e., $\exp(x)\approx 1+x$ for $x \ll 1$) indicates that the process zone length is independent of the crack length and equals L_c. This is the realm of small-scale yielding, $R=\text{const}=L_c \ll L$, $F \approx L$, in which cracks propagate by preserving the structure of the near-tip stress field. By comparing eqns [37] and [3], one can see that the general scaling relationship suggested by dimensional arguments holds; for the case of a constant breakdown stress σ_s within the process zone, the nondimensional coefficient γ in estimate [3] equals $\pi/4(1-\nu)$. Exact analytic solutions assuming small-scale yielding give rise to $\gamma=\pi/2(1-\nu)$ for a linear slip-weakening relationship (see **Figure 3**) (e.g., Chen and Knopoff, 1986) and $\gamma=9\pi/16(1-\nu)$ for a breakdown stress that linearly decreases with distance away from the crack tip (Palmer and Rice, 1973). In the microcrack limit, the length of the equilibrium process zone is not constant even if the critical weakening displacement D_c and the strength drop ($\sigma_s-\sigma_d$) are intrinsic material properties

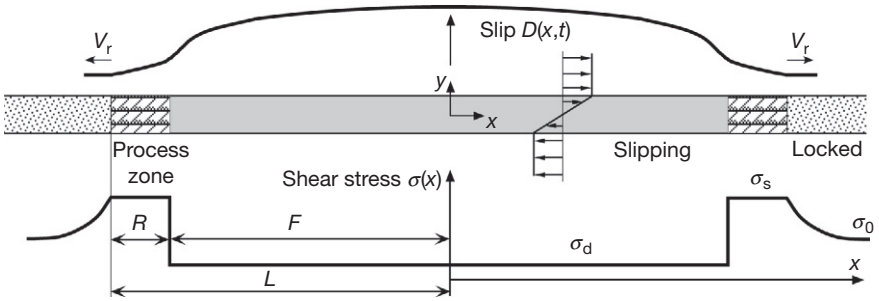


Figure 4 A two-dimensional plain strain shear crack in an infinite elastic medium. The imposed shear stress at the infinity is σ_0. The developed part of the crack is assumed to have a constant residual shear stress σ_d. At the crack tips, there are process zones having length R and shear stress σ_s.

independent of the ambient stress and loading conditions. In particular, R is predicted to exponentially increase as the length of the stress drop region F decreases. Equations [38] and [32] suggest that the presence of sufficiently short cracks (such that $F<L_c$) has no effect on the macroscopic 'strength' of rocks. In particular, the prestress required for the crack extension must approach the static yield limit, $\sigma_0 \to \sigma_s$, and the size of the yield zone increases without bound, $L \approx R \to \infty$, for $F \to 0$ (see eqn [38]).

Quantitative estimates of the critical length scale L_c are not straightforward because it is not clear whether the slip-weakening distance D_c is indeed a scale-independent material constant (e.g., Barton, 1971; Ohnaka, 2003; Rudnicki, 1980). Laboratory measurements of the evolution of friction on smooth slip interfaces indicate that D_c may be of the order of 10^{-5} m (Dieterich, 1979; Li, 1987; Marone, 1998). For $\mu \sim 10^{10}$ Pa and $(\sigma_s - \sigma_d) \sim 10^7$–$10^8$ Pa (likely spanning the range of strength drops for both 'strong' and 'weak' faults), from eqn [36], one obtains $L_c \sim 10^{-3}$–10^{-1} m, negligible compared to the characteristic dimension of the smallest recorded earthquakes but comparable to the typical sample size used in the laboratory experiments. An upper bound on L_c may be obtained from estimates of the effective fracture energies of earthquake ruptures. For large (moment magnitude >6) earthquakes, the seismically inferred fracture energies $G_c = (\sigma_s - \sigma_d) D_c$ are of the order of 10^6–10^7 J m^{-2} (Abercrombie and Rice, 2005; Beroza and Spudich, 1988; Husseini, 1977; Ida, 1972), rendering the effective $D_c \sim 0.01$–1 m. Assuming that the seismically inferred values of D_c are applicable to quasistatic cracks, eqn [36] suggests that a transition from 'micro' to 'macro' rupture regimes occurs at length scales of the order of 1–10^3 m.

The magnitude of a prestress required to initiate the crack propagation can be found by combining eqns [37] and [32]:

$$\frac{\sigma_s - \sigma_0}{\sigma_s - \sigma_d} = \frac{2}{\pi} \arcsin \exp\left(-\frac{L_c}{F}\right) \qquad [39]$$

In the microcrack or large-scale yielding limit, $F \ll L_c$, eqn [39] predicts $\sigma_0 \approx \sigma_s$, that is, the crack propagation requires ambient stress comparable to the peak static strength of crustal rocks, as discussed in the preceding text. In the small-scale yielding limit, $F \gg L_c$, eqn [39] gives rise to a well-known inverse proportionality between the stress drop $(\sigma_0 - \sigma_d)$ and the square root of the crack length F (e.g., Cowie and Scholz, 1992; Kostrov, 1970; Rice, 1968a):

$$\sigma_0 = \sigma_d + \frac{2}{\pi}(\sigma_s - \sigma_d)\left(\frac{2L_c}{F}\right)^{1/2} \qquad [40]$$

It follows from eqn [40] that for sufficiently large ruptures, the background tectonic stress required for the rupture propagation does not need to appreciably exceed the residual friction on the slipped surface σ_d. That is, the stress drop associated with the rupture propagation, $(\sigma_0 - \sigma_d)$, may be much smaller than the strength drop, $(\sigma_s - \sigma_d)$, provided that the rupture size significantly exceeds the critical nucleation size L_c. This statement forms the basis of the 'statically strong, but dynamically weak' fault theory (Lapusta and Rice, 2004). According to this theory, major crustal faults may operate at relatively low driving stresses (e.g., sufficient to explain the so-called heat flow paradox of the San Andreas fault (Brune et al., 1969; Lachenbruch, 1980)), even if the peak failure stress required for the onset of dynamic weakening is consistent with Byerlee's law and hydrostatic pore pressures (Byerlee, 1978; Marone, 1998; Scholz, 2002), provided that $\sigma_d \ll \sigma_s$. If so, earthquake ruptures must nucleate in areas where σ_0 approaches σ_s (due to either locally increased ambient stress or decreased static strength, e.g., due to high pore fluid pressures) and propagate into areas of relatively low ambient stress. Under this scenario, the overall fault operation must be such that the average stress drop $\Delta\sigma$ remains relatively small (of the order of 0.1–10 MPa) and essentially independent of the rupture size (Abercrombie, 1995; Kanamori and Anderson, 1975; Scholz, 2002). Because the overall seismic moment release is dominated by the largest events, the implication from the Lapusta and Rice (2004) model is that the Earth crust is not able to support high deviatoric stresses in the vicinity of large active faults. Phenomenologically, this is consistent with the 'weak fault' theory maintaining that the average shear stress σ_0 resolved on mature faults is of the order of the earthquake stress drops (i.e., up to a few tens of megapascals) and is considerably less than predictions based on Byerlee's law (a few hundreds of megapascals) (e.g., Kanamori and Heaton, 2000). The 'statically strong but dynamically weak' fault theory seeks to reconcile laboratory results from rock friction experiments with seismic observations. Neither the peak shear stress σ_s nor the residual dynamic friction σ_d can be estimated from seismic data. Both parameters are likely scale-dependent; for example, the peak shear stress σ_s may vary from gigapascals on the scale of microasperities and gouge particles (10^{-6} m) to the Mohr–Coulomb stress $f_s(\sigma_n - p)$, where f_s is the static coefficient of friction, σ_n is the fault-normal stress, and p is the pore fluid pressure, on the scale of centimeters to meters (i.e., consistent with laboratory data), to values that may be lower still on scales of hundreds of meters to kilometers. Similarly, the residual dynamic friction may depend on the amount of slip (and the rupture size) (e.g., Abercrombie and Rice, 2005; Brown and Fialko, 2012; Kanamori and Heaton, 2000; Rice, 2006); implications from such behavior are further discussed in **Section 4.03.7**.

4.03.5 Shear Cracks Governed by Rate-and-State Friction Laws

More sophisticated models of earthquake faults combine elasticity with empirical friction laws that relate shear stress to normal stress and slip velocity and slip history. In particular, the rate-and-state phenomenology (Dieterich, 1979; Ruina, 1983) provides a system of differential equations that can be used to define shear stress (e.g., σ in eqn [30]) on a fault plane without oversimplifying assumptions, such as eqns [32]. The resulting system of equations precludes analytic solutions, and numerical experiments are required to investigate the evolution of stress and slip rate on a fault. In the framework of rate-and-state friction, material properties that govern fault slip are a (rate parameter), b (state parameter), and the characteristic evolution distance analogous to the critical weakening displacement D_c discussed in the preceding text (e.g., **Figure 3**; also, see **Chapter 4.04**).

In particular, numerical simulations performed by Dieterich (1992) revealed that slip instabilities on a

rate-and-state fault require a nucleation zone having a characteristic size greater than $L_b = D_c\mu/[(1-\nu)b\sigma_n]$, where σ_n is the fault-normal stress (see Figure 4 in **Chapter 4.04**). The existence of a minimum nucleation length scale was previously suggested based on a conditional stability of a spring-box slider (e.g., Gu et al., 1984; Rice, 1993; Ruina, 1983). However, the spring-box slider model predicts that the minimum nucleation length should scale as $(b-a)^{-1}$, rather than b^{-1}, as reported by Dieterich (1992). This discrepancy may stem from the fact that slip on a deformable fault is intrinsically a higher dimensional (2-D or 3-D) problem compared with a (1-D) problem of motion of a spring-box slider.

Rubin and Ampuero (2005) showed that L_b is in fact the minimum nucleation length and that for a weakly weakening ($a/b \to 1$, that is, nearly velocity neutral) friction, larger nucleation areas can exist, with the characteristic length of the order of $b^2/(b-a)^2 L_b$. The latter can be reached under quasi-steady-state conditions (such that the state variable is nearly constant), while L_b is the minimum nucleation length when the fault is well above steady state (the state variable is rapidly changing) (Ampuero and Rubin, 2008; Rubin and Ampuero, 2005).

Some parallels may be drawn between the analytic solutions presented in **Section 4.03.4** and numerical results for rate–state faults (Ampuero and Rubin, 2008; Dieterich, 1992; Rubin and Ampuero, 2005). If one approximates the velocity change at the tip of a slipping patch on a rate–state fault by a step function (e.g., see Figure 8 in **Chapter 4.02**), then parameters $a\sigma_n$ and $b\sigma_n$ of rate-and-state friction are analogous to the difference between the yield strength and the background stress ($\sigma_s - \sigma_0$) and the strength drop ($\sigma_s - \sigma_d$), respectively (see **Figure 4** in this chapter and Figure 1 in **Chapter 4.04**). The characteristic length of the process zone L_c (eqn [36]) is then essentially coincident with the minimum nucleation size L_b established by Dieterich (1992) based on numerical simulations coupling the rate-and-state friction and elasticity. The correspondence between the two length scales may be interpreted as indicating that no elastodynamic instability is possible in the microcrack regime ($F < L_c$). On the other hand, a condition $a/b \to 1$ implies $\sigma_0 \approx \sigma_d$, that is, small-scale yielding (see eqn [40]). Indeed, such conditions were satisfied prior to the onset of elastodynamic instability in numerical experiments of Rubin and Ampuero (2005) performed under the assumption of weak velocity weakening. Using eqns [32] and [38], one can see that in the limit $R/L \ll 1$, $b^2/(b-a)^2 L_c \sim L/R$ $L_c = L$, that is, the 'velocity neutral' nucleation length of Rubin and Ampuero (2005) is essentially the length of a quasistatic crack on the verge of propagation under a constant remote stress σ_0. This is the maximum length of a slip patch satisfying conditions of a quasistatic equilibrium. The main differences between the rate–state model and that shown in **Figure 4** is that the latter assumes a piece-wise constant stress distribution on the crack surface and zero slip velocity ahead of the crack tip, while in the rate–state model, the shear stress is a continuous function of distance along the fault, and the fault slip rate is never zero. This analogy between the analytic solution for a static crack and a deformable rate–state fault may hold because of large (orders of magnitude) increases in the slip velocity at the tip of an accelerating slip patch compared to the nominally locked region ahead of the tip; variations in slip rate behind the rupture front are much smaller in comparison, so that a steplike change in slip rate may be a reasonable approximation.

4.03.6 Dynamic Effects

Quasistatic solutions considered in **Section 4.03.4** are valid only for rupture speeds that are well below the shear wave velocity V_s. As the rupture velocity increases, the inertial term in the equilibrium equations [10] eventually becomes nonnegligible, and the near-tip stress field is significantly altered when V_r becomes a sizeable fraction of V_s (Andrews, 1976a; Broberg, 1978; Freund, 1979; Rice, 1980). The most pronounced effects of a high rupture speed are the relativistic shrinking of the in-plane process zone R and simultaneous increase of stress perturbations off the crack plane (Kame and Yamashita, 1999; Poliakov et al., 2002; Rice et al., 2005). The net result is an increased tendency for branching, bifurcation, and nonsteady propagation, all of which significantly complicate the analytic and numerical treatment of the elastodynamic rupture problem. In particular, the intermittent propagation of the rupture front invalidates the equivalence between the LEFM and slip-weakening formulations (e.g., Freund, 1979), implying a greater dependence of the model results on a specific choice of fracture criteria.

Analysis of the full elastodynamic equilibrium equations [10] reveals that solutions exist for rupture velocities below a limiting speed V_l, which equals to the Rayleigh wave velocity V_R for mode II cracks and shear wave velocity for mode III cracks (e.g., Freund, 1979; Kostrov and Das, 1988). Solutions become singular as $V_r \to V_l$; in particular, the dynamic stress intensity factor and the energy release rate at the crack tip asymptotically vanish. Main fracture mechanics parameters of a steady-state elastodynamic rupture (such as the process zone size and the stress intensity factors) may be readily obtained by multiplying or dividing the respective results from quasistatic solutions by dimensionless coefficients that depend on rupture velocity only. The corresponding coefficients are

$$f_{II} = (1-\nu)\frac{4\sqrt{1-\frac{V_r^2}{V_p^2}}\sqrt{1-\frac{V_r^2}{V_s^2}} - \left(2-\frac{V_r^2}{V_s^2}\right)^2}{\frac{V_r}{V_s}\sqrt{1-\frac{V_r^2}{V_s^2}}} \qquad [41]$$

$$f_{III} = \sqrt{1-\frac{V_r^2}{V_s^2}} \qquad [42]$$

for the mode II and mode III loading, respectively (e.g., Freund, 1998; Rice, 1980). In eqn [41], V_p is the P-wave velocity. Coefficients f_{II} and f_{III} monotonically decrease from unity at $V_r = 0$ to zero at $V_r = V_l$. For instance, the length of the dynamic process zone R^* is given by $R^*(V_r) = f_{II,III}(V_r)R$, and the dynamic stress intensity factor for a self-similar expanding crack is $K_{II,III}(V_r) = f_{II,III}(V_r)K_{II,III}(0)$. Note that for a steady-state self-healing pulse with a fixed stress distribution in the reference frame of a moving pulse, the process zone shortens at high rupture velocity, similar to the case of a self-similar crack, but the dynamic stress intensity factor is independent of V_r.

Rice et al. (2005) confirmed these results with full analytic solutions for a self-healing pulse with a linearly weakening process zone, propagating at a constant rupture speed. They found that the ratio of the dynamic process zone R^* to the pulse length L^* is independent of the rupture speed but is dependent on the ratio of stress drop to strength drop:

$$\frac{\sigma_0 - \sigma_d}{\sigma_s - \sigma_d} = \frac{\xi}{\pi} - \frac{\xi - \sin\xi}{2\pi \sin^2(\xi/2)} \qquad [43]$$

where $\xi = 2\arcsin\sqrt{R^*/L^*}$ (cf. eqn [32]). The velocity invariance of the relative size of the process zone gives rise to a somewhat unintuitive result that the pulse length L^* vanishes as the rupture accelerates to a limiting speed. At the same time, the amount of slip produced by a self-healing pulse is also invariant with respect to the rupture velocity. Thus, the dynamically shrinking rupture size gives rise to a dramatic increase in the near-field coseismic strain. Figure 5 shows the dependence of the near-tip stress concentration on the rupture speed inferred from the Rice et al. (2005) model. As one can see from Figure 5, the extent of the off-plane yielding and damage substantially increases at high rupture speeds, so that the assumption of a thin in-plane process zone ceases to be valid. Similar results were reported for a semi-infinite crack by Poliakov et al. (2002). Note that the deduced areas of yielding likely underestimate the extent of off-fault damage, as stresses inside the shaded areas in Figure 5 are beyond the failure envelope. By explicitly allowing inelastic deformation, the excess stresses will be relaxed, and the yielding zone will further expand. The enhanced off-fault damage may be one of the mechanisms preventing extreme contractions of the rupture length at high propagation velocities. It may also appreciably modify the expenditure part of the earthquake energy balance, as discussed in the following section.

Elastodynamic slip instabilities with rupture velocity below the limiting speed V_l are referred to as subsonic or subshear rupture. The majority of earthquakes for which high-quality measurements of the rupture speed are available appear to be subsonic. Theoretical models indicate that immediately above the limiting rupture speed V_l, the flux of mechanical energy into the crack tip region becomes negative for tensile and mode II shear cracks, effectively prohibiting self-sustained fracture (Broberg, 1999; Freund, 1998). However, physically admissible solutions do exist for shear cracks with rupture speeds spanning the interval between the S-wave and P-wave velocities; such mode of propagation is referred to as intersonic, transonic, or supershear rupture (Andrews, 1976b; Broberg, 1989; Burridge, 1973; Simonov, 1983). Although the mechanics of transition from the subsonic ($V_r < V_l$) to transonic

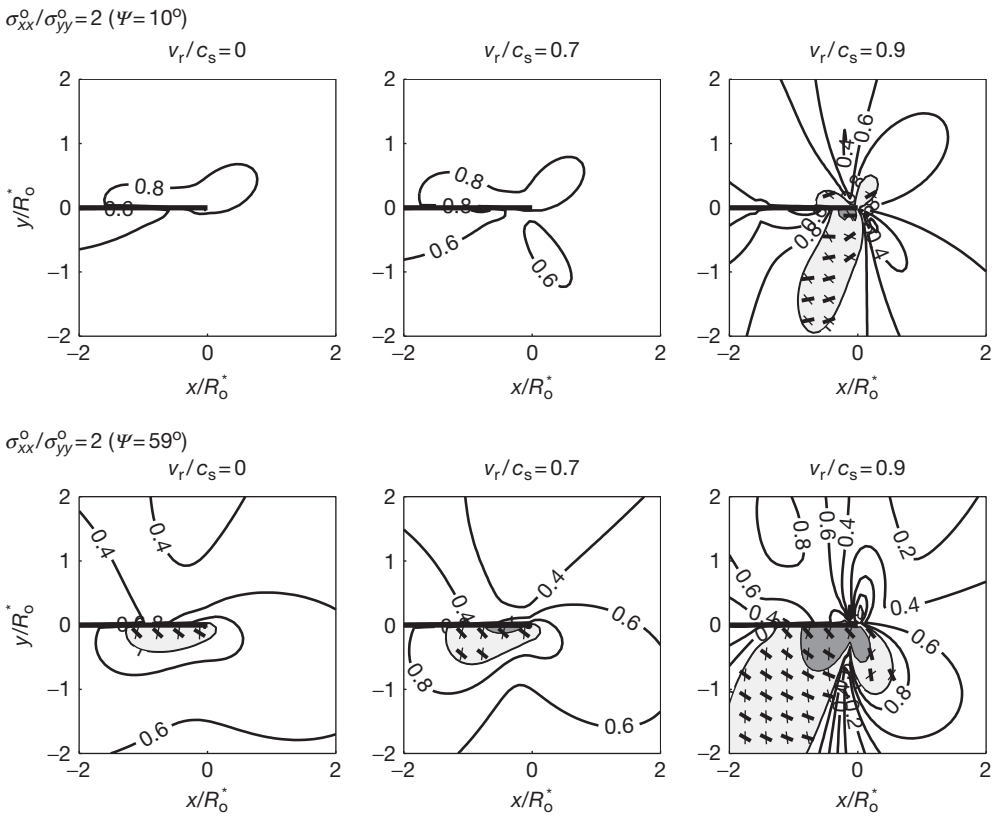


Figure 5 Stress perturbation around the tip of a propagating slip pulse with a linear slip-weakening process zone (reproduced from Rice JR, Sammis CG, and Parsons R (2005) Off-fault secondary failure induced by a dynamic slip pulse. *Bulletin of the Seismological Society of America* 95: 109–134, their Figure 6). Light shading denotes areas where the Mohr–Coulomb failure envelope is exceeded (likely to fail in shear), and dark shading denotes areas of absolute tension (likely to fail by tensile cracking). Thick tick marks indicate the orientation of optimal planes for right-lateral slip, and thin tick marks indicate the orientation of optimal planes for left-lateral slip. Ψ is an angle between the maximum compression axis and the fault plane. Process zone comprises 10% of the rupture length. Axes are normalized by the quasistatic process zone size.

($V_r > V_s$) propagation is not fully understood, there is experimental (Samudrala et al., 2002; Xia et al., 2004) and seismologic (Archuleta, 1984; Bouchon and Vallee, 2003; Dunham and Archuleta, 2004; Wald and Heaton, 1994) evidence that transonic rupture speeds may be achieved under certain conditions. The spatial structure of the near-tip stress field and the radiation pattern in the transonic regime are markedly different from those due to subsonic ruptures (see **Chapter 4.08**).

4.03.7 Fracture Energy

The concept of fracture energy was originally introduced for tensile cracks by Griffith (1920) to quantify the irreversible work associated with breaking of the intermolecular bonds and creation of a stress-free crack surface. Griffith's definition based on a global energy balance was subsequently shown to be equivalent to local definitions based on the LEFM and small-scale yielding models (Rice, 1968a; Willis, 1967). For example, for a Barenblatt-type process zone model, the fracture energy is the work spent against the cohesive stress σ_s in the process zone on separating the crack walls by the critical opening distance D_c. An elegant demonstration of the equivalence of global and local definitions of fracture energy for tensile cracks was provided by Rice (1968b) in a form of the path-independent *J*-integral (also, see Cherepanov, 1968; Eshelby, 1956). Palmer and Rice (1973) extended this technique to the case of shear cracks and defined the shear fracture energy as work required to evolve shear stress on the slip interface from the yield stress (or static friction) σ_s to the residual dynamic friction σ_d:

$$G_c = \int_0^{D_c} (\sigma(D) - \sigma_d) \mathrm{d}D \qquad [44]$$

where $\sigma(D)$ varies between σ_s and σ_d for $0 < D < D_c$, respectively. A similar formulation was introduced by Ida (1972). Equation [44] allows a simple insight into the fracture process and has been widely used for interpretations of seismic data. However, several factors may limit its application to the analysis of earthquake ruptures. First, the displacement-weakening model assumes that all inelastic deformation is limited to the slip plane. Both theoretical models (Andrews, 2005; Rice et al., 2005; Rudnicki, 1980) (**Figure 5**) and field observations (Chester et al., 2005; Fialko, 2004b; Fialko et al., 2002; Li et al., 1998) suggest that the earthquake-induced damage likely extends well off of the fault plane, and the energy dissipated in the fault damage zone may be quite significant (e.g., Andrews, 2005; Ben-Zion and Shi, 2005; Wilson et al., 2004). Second, the fracture energy given by eqn [44] has a clear physical interpretation if the residual dynamic stress σ_d is constant (or at least if the along-fault variations in σ_d are small compared to the strength drop, $\sigma_s - \sigma_d$). The second point can be illustrated by considering a traditional representation of the earthquake energy budget:

$$\Delta U_p = U_r + U_f + U_G \qquad [45]$$

where ΔU_p is the change in the total potential energy (which includes changes in the elastic strain energy ΔU_e and gravitational potential energy); U_r is the energy radiated in seismic waves; U_f is the energy dissipated on the well-slipped portion of the fault due to friction, comminution, phase transitions, and other irreversible losses; and U_G is the fracture energy spent on overcoming high resisting stresses near the crack tip (Dahlen, 1977; Kostrov, 1974; Rivera and Kanamori, 2005; Rudnicki and Freund, 1981). A significant part of U_f is believed to be ultimately converted into heat (Fialko, 2004a; Sibson, 1980). Under the approximation of the displacement-weakening model, $U_G \propto \sigma_s D_c$, and $U_f \propto \sigma_d D_m$. Assuming that the residual friction is of the order of the peak strength, $\sigma_d \sim O(\sigma_s)$, and the critical slip-weakening distance is much smaller than the coseismic offset, $D_c \ll D_m$, the fracture energy is negligible compared to frictional losses in the earthquake energy balance equation [45]. However, if the fault friction progressively decreases with slip, as suggested by the experimental observations and theoretical inferences of the dynamic weakening (Abercrombie and Rice, 2005; Di Toro et al., 2004; Fialko and Khazan, 2005; Goldsby and Tullis, 2002; Tsutsumi and Shimamoto, 1997), the effective slip-weakening distance D_c is expected to scale with the slip magnitude, and the fracture energy U_G may not be small compared to U_f. Because neither the slip-weakening distance D_c nor the residual friction σ_d in this case are material properties (in particular, they may depend on the details of slip history, thickness and permeability of the slip zone, etc.), a distinction between U_G and U_f terms in the earthquake energy balance equation [45] becomes somewhat arbitrary. Note that for a rupture on a preexisting fault, there is little *physical* difference between U_G and U_f, as both terms represent spatially and temporally variable frictional losses associated with fault slip; both U_G and U_f ultimately contribute to wearing and heating on the slip interface. The situation is further complicated if the dynamic friction is a nonmonotonic function of slip (e.g., Brune and Thatcher, 2002; Hirose and Shimamoto, 2005; Rivera and Kanamori, 2005; Tinti et al., 2005). While a formal distinction between the frictional and fracture losses associated with shear ruptures may be problematic, the entire amount of work spent on inelastic deformation of the host rocks during the crack propagation is unambiguous and can be readily quantified. For simplicity, here, we consider the case of a quasistatic crack growth. Formulation presented in the succeeding text can be also generalized to the case of dynamic cracks.

Consider an equilibrium mode II crack in a medium subject to initial stress σ_{ij}^0. The medium is elastic everywhere except inside the crack and within a finite process zone near the crack tips (**Figure 6(a)**). The inelastic zone is demarcated by a surface S_f. Let external forces do some work δW on a medium, in the

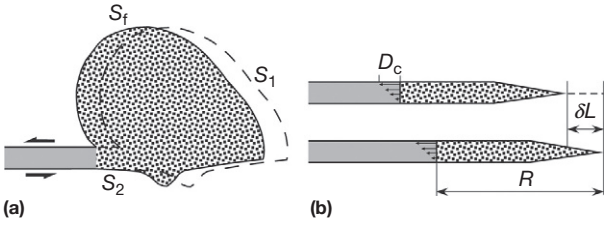


Figure 6 Schematic view of inelastic deformation associated with crack propagation. (a) Finite damage zone extending off the fault plane. (b) Thin in-plane damage zone.

result of which the crack acquires a new equilibrium configuration. In the new configuration, some area ahead of the crack front undergoes inelastic yielding and joins the process zone (see an area bounded by surfaces ΔS_1 and S_f in Figure 6(a)). At the trailing end of the process zone, slip exceeds D_c, and some fraction of the process zone (bounded by surface ΔS_2 in Figure 6(a)) joins the developed part of the crack. The external work δW is spent on changes in the elastic strain energy δU_e and irreversible inelastic deformation δU_G (which includes friction, breakdown, and comminution):

$$\delta W = \delta U_e + \delta U_G \qquad [46]$$

Changes in the elastic strain energy are given by (e.g., Landau and Lifshitz, 1986; Timoshenko and Goodier, 1970)

$$\delta U_e = \frac{1}{2}\left(\sigma_{ij}^1 \epsilon_{ij}^1 - \sigma_{ij}^0 \epsilon_{ij}^0\right) \qquad [47]$$

where σ_{ij}^1 and ϵ_{ij}^1 are stresses and strains, respectively, in the elastic part of a medium after the crack extension. The assumption of linear elasticity [9] implies that

$$\sigma_{ij}^0 \epsilon_{ij}^1 \equiv \sigma_{ij}^1 \epsilon_{ij}^0 \qquad [48]$$

for any σ_{ij}^k and ϵ_{ij}^k. The identity eqn [48] allows one to write eqn [47] as follows:

$$\delta U_e = \frac{1}{2}\left(\sigma_{ij}^1 + \sigma_{ij}^0\right)\epsilon_{ij}^1 - \frac{1}{2}\left(\sigma_{ij}^1 + \sigma_{ij}^0\right)\epsilon_{ij}^0 = \frac{1}{2}\left[\left(\sigma_{ij}^1 + \sigma_{ij}^0\right)\delta u_i\right]_{,j} \qquad [49]$$

where $\delta u_i = u_i^1 - u_i^0$ is the displacement field produced by crack propagation. Expressions [47] and [49] assume that the strains are infinitesimal, so that the relationship between strain and displacement gradients is given by eqn [8]. Also, expression [49] makes use of the fact that under quasistatic conditions, the divergence of stress is zero, $\sigma_{ij,j} = 0$ (see the equilibrium equations [10]; note that the body forces may be excluded from consideration by incorporating the effects of gravity in prestress). Using the Gauss theorem along with a condition that the crack-induced deformation must vanish at infinity, from the energy balance equation [46], one obtains the following expression for the work done on inelastic deformation:

$$\begin{aligned}\delta U_G &= \delta W - \delta U_e \\ &= -\int_{S_f} \sigma_{ij}^0 n_j \delta u_i \mathrm{d}S - \frac{1}{2}\int_{\Delta S_1 + \Delta S_2} \left(\sigma_{ij}^0 + \sigma_{ij}^1\right) n_j \delta u_i \mathrm{d}S\end{aligned} \qquad [50]$$

In the limit of an ideally brittle fracture, the area of inelastic yielding has a negligible volume (i.e., $S_f \to 0$), so that the first integral on the right-hand side of eqn [50] vanishes. In the second integral, the surface ΔS_2 also vanishes, while the surface ΔS_1 becomes the crack length increment δL (Figure 6(b)). Within δL, the stresses are weakly singular, $\sigma_{ij} \propto 1/\sqrt{r}$, where r is distance to the crack tip, and the crack wall displacements scale as $\delta u_i \propto \sqrt{r}$, so that the product of stresses and displacements is of the order of unity, and the corresponding integral in eqn [50] is of the order of $\Delta S_1 = \delta L$. This is the well-known LEFM limit, for which the fracture energy is given by eqn [2].

For more realistic models that explicitly consider failure at the crack tip, the stresses are finite everywhere, so that the second integral on the right-hand side of eqn [50] is of the order of $\sigma_{ij}^0 \delta u_i (\Delta S_1 + \Delta S_2)$, that is, negligible compared to the integral over the finite inelastic zone $\sim O(\sigma_{ij}^0 \delta u_i S_f)$. Provided that the displacement field associated with the crack extension can be represented as $\delta u_i = \delta L \partial u_i / \partial L$, eqn [50] allows one to introduce the fracture energy G_c as the total inelastic work per increment of the crack length or the energy release rate:

$$G_c = \frac{\partial U_G}{\partial L} = -\frac{1}{2}\int_{S_f} \frac{\partial u_i}{\partial L} \sigma_{ij}^0 n_j \mathrm{d}S \qquad [51]$$

Factor of 1/2 in eqn [51] stems from the assumption of a bilateral crack propagation. Equation [51] in general cannot be readily evaluated analytically because the size and geometry of the inelastic zone S_f are not known in advance and have to be found as part of a solution. Further insights are possible for special cases. For example, assuming that all yielding is confined to a crack plane (Figure 6(b)), eqn [51] reduces to

$$G_c = \int_0^L \sigma(x) \frac{\partial D(x)}{\partial L} \mathrm{d}x \qquad [52]$$

where we took into account that the total offset between the crack walls is $D(x) = 2u_x$, and the sense of slip is opposite to that of the resisting shear tractions acting on the crack walls. The integral equation [52] is still intractable for an arbitrary loading, as the derivative $\partial D(x)/\partial L$ must be calculated along the equilibrium curve (e.g., see eqn [33]). Closed form analytic solutions can be obtained for limiting cases of small-scale ($L \gg R$) and large-scale ($L \approx R$) yielding. First, consider a crack that is much longer than the critical size L_c [36] and has a complete stress drop, $\sigma(x) = 0$ for $D > D_c$. As shown in Section 4.03.2, for such a crack the size of the process zone is independent of the crack length, $R = L_c$, and the crack propagation does not modify the slip distribution within the process zone in the reference frame of a propagating crack tip. In this case,

$$\partial D(x)/\partial L = -\partial D(x)/\partial x \qquad [53]$$

so that eqn [52] gives rise to

$$G_c = \int_{L-R}^{L} \sigma(x) \frac{\partial D(x)}{\partial L} \mathrm{d}x = \int_0^{D_c} \sigma(D) \mathrm{d}D \qquad [54]$$

Expression [54] is analogous to the result obtained using the J-integral technique (Rice, 1968b). An expression for the J-integral contains a derivative of the crack wall displacement with respect to the integration variable, rather than the crack length. As noted by Khazan and Fialko (2001), the two derivatives coincide (up to a sign) in a limiting case of a very long crack (for which the J-integral was derived), but the difference may be significant if the small-scale yielding approximation does not hold (also, see Rice, 1979).

For a case of a constant, but nonvanishing, residual friction, $\sigma(x) = \sigma_d$ for $D_c < D < D_m$, evaluation of integral equation [52] gives rise to

$$G_c = \sigma_d (D_m - D_c) + \int_0^{D_c} \sigma(D) \mathrm{d}D \qquad [55]$$

thanks to self-similarity of the along-crack displacement $D(x)$ over the interval $0 < x < L - R$ for long cracks, such that the

relationship [53] still holds. For a constant yield stress within the process zone, $\sigma(x) = \text{const} = \sigma_s$ (e.g., **Figure 4**), eqn [55] gives rise to a simple expression $G_c = (\sigma_s - \sigma_d)D_c + \sigma_d D_m$, in which the first and second terms may be recognized as the traditionally defined fracture energy U_G and the frictional work U_f, respectively (cf. eqn [45]).

In case of a large-scale yielding, relationship [53] is generally not applicable. Assuming $\sigma_s = \text{const}$, expression [52] can be integrated by parts to yield

$$G_c = \sigma_s \frac{\partial}{\partial L}\int_0^L D(x)\mathrm{d}x - \sigma_s \int_0^F \frac{\partial D}{\partial L}\mathrm{d}x \qquad [56]$$

In the developed part of the crack with the full stress drop ($x < F$), slip is essentially constant and equals to D_c, so that the second integral on the right-hand side of eqn [56] can be neglected. Taking advantage of expressions [33] and [35] to evaluate the first term in eqn [56], one obtains (Khazan and Fialko, 2001)

$$G_c = \frac{\pi F}{2L_c}\sigma_s D_c \qquad [57]$$

Equation [57] indicates that the fracture energy for the case of small cracks (or large-scale yielding, $F \ll L_c$) is substantially different from the fracture energy for large cracks (or small-scale yielding, $F \gg L_c$). In case of large-scale yielding, the fracture energy is not constant even if both the yield strength σ_s and the critical slip-weakening displacement D_c are material constants independent of loading conditions. In particular, G_c is predicted to linearly increase with the size of the developed part of the crack F. A linear scaling of fracture energy implies increases in the apparent fracture toughness K_q proportional to a square root of the developed crack length, $K_Q \propto \sqrt{F}$ (see eqn [2]), in the large-scale yielding regime. Increases in the apparent fracture toughness for small cracks are well known from laboratory studies of tensile fracture (e.g., Bazant and Planas, 1998; Ingraffea and Schmidt, 1978).

The apparent scaling of the earthquake fracture energy with the earthquake size has been inferred from seismic data (e.g., Abercrombie and Rice, 2005; Husseini, 1977; Kanamori and Heaton, 2000). Arguments presented in the preceding text indicate that several mechanisms may be responsible for the observed increases in the seismically inferred fracture energies with the rupture length, in particular, (1) off-fault damage that scales with the rupture length (larger ruptures are expected to produce broader zones of high stress near the rupture fronts, presumably advancing the extent of off-fault damage (see **Figure 5** and eqn [51])); (2) a continuous degradation of dynamic friction on a fault plane, for example, due to thermal pressurization or any other slip-weakening mechanism; and (3) rupture propagation under conditions of large-scale yielding (eqn [57]), although it remains to be seen whether elastodynamic instability can occur when the process zone comprises a substantial fraction of the crack length (see **Section 4.03.4**). These mechanisms are not mutually exclusive and may jointly contribute to the observed scaling $G_c \propto F$. For example, the third mechanism might be relevant for small earthquakes, while the first and second ones perhaps dominate for large events. Note that the second mechanism is ultimately limited by a complete stress drop, beyond which no further increase in fracture energy is possible. The same limit may also apply to the first mechanism, as the size of the dynamic damage zone scales with the quasistatic one for a given displacement-weakening relationship (Rice et al., 2005). Establishing the relative importance of contributions of various mechanisms to the effective fracture energy is an important but challenging task. In particular, if contributions from the off-fault yielding are substantial, interpretations of seismic data that neglect such yielding may systematically overestimate the magnitude of the effective slip-weakening distance D_c. Unfortunately, distinguishing between different contributions to the overall value of G_c can be unlikely accomplished based on the seismic data alone.

In summary, the fracture energy defined by eqn [51] is analogous to Griffith's concept for tensile cracks, provided that the stress drop is complete ($\sigma_d = 0$) and the small-scale yielding condition is met. For a nonvanishing friction on the crack surface, eqn [51] combines inelastic work spent against residual friction and work spent on evolving the shear stress on a fault to a residual level (i.e., the traditionally defined fracture energy). Separation between these two contributions is justified if small-scale yielding condition applies but may be ill-defined otherwise. For models with a continuous strength degradation, there is a continuous repartitioning of the energy budget, such that the effective fracture energy increases at the expense of a diminishing frictional dissipation.

4.03.8 Coupling between Elastodynamics and Shear Heating

One of the factors that can strongly affect the dynamic friction on the slipping interface and, thereby, the seismic radiation, efficiency, and stress drop is the coseismic frictional heating. Rapid slip during seismic instabilities may substantially raise temperature on a fault surface. The dependence of dynamic friction on temperature may stem from several mechanisms, including thermal pressurization by pore fluids (Lachenbruch, 1980; Mase and Smith, 1987; Sibson, 1973), frictional melting (Jeffreys, 1942; Maddock, 1986; McKenzie and Brune, 1972; Sibson, 1975), and flash heating of contact asperities (Rice, 2006). Recent experimental measurements confirm appreciable variations in the dynamic friction at slip velocities approaching the seismic range of order of a meter per second (Brown and Fialko, 2012; Di Toro et al., 2004; Goldsby and Tullis, 2002, 2011; Hirose and Shimamoto, 2003; Reches and Lockner, 2010; Spray, 2005). The documented variations are significantly larger than predictions of the rate-and-state friction (Dieterich, 1979; Ruina, 1983) extrapolated to seismic slip rates. The likely importance of the thermally induced variations in friction warrants a quantitative insight into the dynamics of fault heating during seismic slip. Major questions include the following: What are the dominant mechanisms of fault friction at high slip rates? How do increases in temperature affect the dynamic shear stress on a slipping interface? How robust is the thermally activated weakening? Is dynamic friction a monotonically decaying function of temperature? If not, what are the mechanisms, conditions, and significance of the thermally activated strengthening? When and where is the onset of the thermally induced weakening or strengthening

likely to occur on a slipping interface? And what are the implications for the dynamics of earthquake ruptures?

In a simple case of the LEFM crack with a constant residual shear stress, the frictional heating problem admits closed analytic solutions (Fialko, 2004a; Richards, 1976). Consider a mode II (plain strain) crack rupturing bilaterally at a constant speed V_r (**Figure 7**), so that $L(t)=tV_r$. The thickness of the gouge layer that undergoes shear, and the shear strain rate across the gouge layer are assumed to be constant (Cardwell et al., 1978; Mair and Marone, 2000). Because the thickness of the gouge layer $2w$ is negligible compared with any other characteristic length scale in the problem (e.g., the rupture size $2L$ and the amount of slip D), the temperature evolution in the gouge layer and in the ambient rock is well described by the one-dimensional diffusion equation with a heat source:

$$\frac{\partial T}{\partial t}=\kappa\frac{\partial^2 T}{\partial y^2}+\frac{Q}{c\rho} \qquad [58]$$

where y is the crack-perpendicular coordinate and Q is the rate of frictional heat generation within the slipping zone,

$$Q(x,y,t)=\begin{cases}\dfrac{\sigma_d(x)}{2w(x)}\dfrac{\partial D(x,t)}{\partial t}, & t>0, |y|<w\\ 0, & |y|>w\end{cases} \qquad [59]$$

$\partial D/\partial t$ being the local slip velocity. A solution to eqn [58] subject to the initial condition $T(x, y, 0)=T_0$, where T_0 is the temperature of the host rocks prior to faulting, is (Fialko, 2004a)

$$T-T_0=\frac{1}{4c\rho w}\int_{x/V_r}^{t}\left(\operatorname{erf}\left[\frac{y+w}{2\sqrt{\kappa(t-\tau)}}\right]-\operatorname{erf}\left[\frac{y-w}{2\sqrt{\kappa(t-\tau)}}\right]\right)\frac{\partial D(x,\tau)}{\partial \tau}\sigma_d(x)\,d\tau \qquad [60]$$

Dimensional arguments suggest the following similarity variables:

$$\text{Nondimensional along-fault coordinate } \chi=\frac{x}{tV_r} \qquad [61]$$

$$\text{Nondimensional fault thickness } \overline{w}=\sqrt{\frac{2}{\kappa t}}w \qquad [62]$$

For the LEFM crack with a constant stress drop, the along-crack displacement profile $D(x, t)$ is self-similar in that it may be expressed in terms of a single similarity variable $\chi=\chi(x, t)$:

$$D(x,t)=L(t)\epsilon\sqrt{1-\chi^2},\quad t>0, |\chi|<1 \qquad [63]$$

where ϵ is the characteristic shear strain due to the crack, $\epsilon=D(0, t)/L(t)=2(1-\nu)(\sigma_0-\sigma_d)/\mu$. Here, ϵ is taken to be independent of L, as the earthquake stress drops $(\sigma_0-\sigma_d)$ do not exhibit any scale dependence across a wide range of earthquake magnitudes (Abercrombie, 1995; Kanamori and Anderson, 1975; Scholz, 2002). The local slip rate in terms of new variables is

$$\frac{\partial D}{\partial t}=\frac{V_r\epsilon}{\sqrt{1-\chi^2}} \qquad [64]$$

Equations [60] and [64] suggest the following similarity variable for temperature:

$$\text{Nondimensional temperature } \theta=\frac{T-T_0}{\hat{T}} \qquad [65]$$

where

$$\hat{T}=\frac{\sigma_d V_r\epsilon}{c\rho}\sqrt{\frac{t}{\pi\kappa}} \qquad [66]$$

is a characteristic temperature scale for frictional heating assuming a perfectly sharp fault contact.

Substituting eqn [64] into eqn [60] and making use of the similarity variables [62] and [65], one obtains the following expression for the along-crack temperature distribution in the middle of the slip zone ($y=0$):

$$\theta(\chi)=\frac{\sqrt{\pi}}{\overline{w}\sqrt{2}}\int_{\chi}^{1}\operatorname{erf}\left[\frac{\overline{w}}{2\sqrt{2(1-\xi)}}\right]\frac{\xi\,d\xi}{\sqrt{\xi^2-\chi^2}} \qquad [67]$$

Solutions to eqn [67] are shown in **Figure 8**. A family of curves in **Figure 8** illustrates a spatiotemporal evolution of temperature on the slipping fault surface. For faults that are

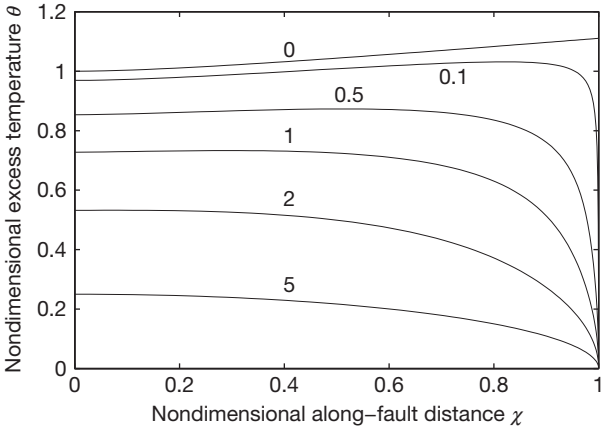


Figure 8 Variations of the nondimensional excess temperature $\theta(\chi)$ along a mode II crack propagating at a constant rupture speed under constant frictional stress. Labels denote the nondimensional thickness of a slipping zone $\overline{w}$ (see eqn [62]).

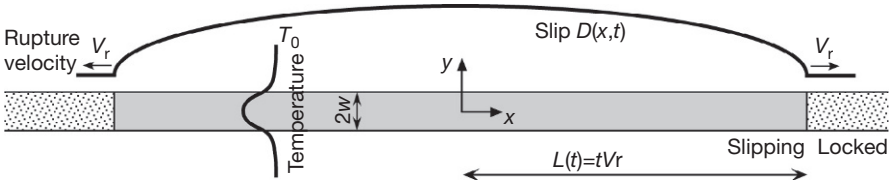


Figure 7 A schematic view of a dynamically propagating mode II crack. The crack has a thickness $2w$ and is rupturing bilaterally at a constant velocity $dL/dt=V_r$.

thicker than the thermal diffusion length scale or at early stages of rupture (i.e., $\overline{w} > 1$), the temperature increase along the fault is proportional to the amount of slip. For thin faults or later during the rupture ($\overline{w} \ll 1$), the temperature is maximum near the crack tip and decreases from the tip to the crack center. However, the instantaneous temperature maximum near the crack tip does not imply cooling of the crack surface behind the tip; at any given point, the temperature on the crack surface $T(x, 0, t)$ steadily increases with time (Fialko, 2004a). For the nondimensional fault thickness $\overline{w}$ of the order of unity, the maximum temperature is reached somewhere between the crack center and the rupture front (**Figure 8**). The inferred anticorrelation between the temperature and the amount of slip stems from a competition between the rates at which the frictional heat is generated at the crack surface and removed to the ambient rocks by conduction. Generation of frictional heat at the tip of the LEFM crack is singular as the thickness of the conductive boundary layer is zero, while the slip velocity is infinite (see eqn [64]). Nonetheless, the excess temperature at the tip is zero for cracks having finite thickness ($\overline{w} > 0$). For cracks that are much thinner than the conductive boundary layer ($\overline{w} \ll 1$), the temperature field develops a shock-like structure, with the tip temperature exceeding the temperature at the crack center by about 10% (**Figure 8**). Assuming that the thickness of the slip zone is constant during an earthquake, eqn [67] predicts that the maximum temperatures are initially attained at the center of a cracklike shear instability. As the earthquake rupture expands, the temperature maximum may migrate toward the rupture fronts. For the thermal diffusivity of the ambient rocks $\kappa = 10^{-6}\ \mathrm{m^2\ s^{-1}}$ and rupture durations of $t = 1$–10 s (corresponding to the rupture sizes of ~5–50 km), this transition will occur for faults that have thickness of the order of $\sqrt{2\kappa t} \sim 2 - 5$ mm or less. The critical fault thickness may be larger still if the heat removal from the fault involves some advective transport by the pressurized pore fluids, and the *in situ* hydraulic diffusivity exceeds the thermal diffusivity κ.

For sufficiently large ruptures, a model of a self-healing slip pulse may be a better approximation than the cracklike models (e.g., Beroza and Mikumo, 1996; Heaton, 1990; Kanamori and Anderson, 1975; Olsen et al., 1997). A self-healing mode II pulse having a constant length L (Freund, 1979) generates a temperature field that is steady state in the reference frame of a moving rupture front (Fialko, 2004a). The appropriate similarity variables are

$$\text{along-fault coordinate } \chi = \frac{x - tV_r}{L} + 1 \qquad [68]$$

$$\text{nondimensional fault thickness } \overline{w} = \sqrt{\frac{2V_r}{L\kappa}}w \qquad [69]$$

$$\text{nondimensional temparature } \theta = \frac{T - T_0}{\hat{T}} \qquad [70]$$

$$\hat{T} = \frac{\sigma_d \epsilon}{c\rho}\sqrt{\frac{LV_r}{\pi\kappa}} \qquad [71]$$

The coseismic displacements and the rate of slip are assumed to have the LEFM-like characteristics at the rupture front ($\chi = 1$) and a nonsingular healing at the trailing edge ($\chi = 0$):

$$D(\chi) = L\epsilon\sqrt{1 - \chi^2} \qquad [72]$$

$$\frac{\partial D}{\partial \chi} = \frac{V_r \epsilon \chi}{\sqrt{1 - \chi^2}} \qquad [73]$$

As before, we assume a constant dynamic friction on the slipping interface. Upon nondimensionalization using variables [71], eqn [60] gives rise to the following expression for the along-fault temperature variations in the middle of the slip zone ($\gamma = 0$):

$$\theta(\chi) = \frac{\sqrt{\pi}}{\overline{w}\sqrt{2}}\int_{\chi}^{1} \mathrm{erf}\left[\frac{\overline{w}}{2\sqrt{2(\chi - \xi)}}\right]\frac{\xi \mathrm{d}\xi}{\sqrt{1 - \xi^2}} \qquad [74]$$

Solutions to eqn [74] are shown in **Figure 9**. The near-tip structure of the temperature field due to a steady-state pulse is similar to that due to a self-similar expanding crack (cf. **Figures 8** and **9**). At the leading edge of an infinitesimally thin shear pulse, there is a thermal shock of amplitude $\hat{T}$ (eqn [71]). The fault temperature monotonically decreases toward the healing front, where the temperature falls to about one-half of the maximum value (**Figure 9**). For 'thick' pulses ($\overline{w} \gg 1$), the fault temperature increases toward the healing front proportionally to the amount of slip. For intermediate nondimensional fault thicknesses of the order of unity, the fault surface initially heats up to a maximum temperature and then cools down before the arrival of the healing front. This behavior is qualitatively different from that on a surface of an expanding crack, which indicates a progressive heating at every point along the crack as long as the rupture continues. The inferred cooling toward the healing front of the steady-state pulse for $\overline{w} > 5$ is caused by a decreasing heat generation due to a vanishing slip velocity and efficient removal of heat by thermal diffusion. For the characteristic rise times L/V_r of the order of seconds, the steady-state LEFM pulses need to be thinner than $5\sqrt{2\kappa L/V_r} \approx 1$ cm to experience maximum temperatures at the rupture front.

Several factors may accentuate the tendency for the temperature peaks near the rupture front. First, higher stresses in the process zone near the rupture tip imply enhanced heating. Numerical simulations indicate that the thermal effect of the

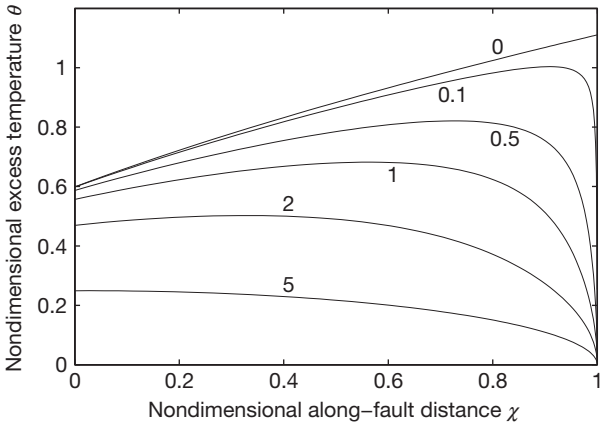


Figure 9 Variations of the nondimensional excess temperature $\theta(\chi)$ along a self-healing pulse. Labels denote the nondimensional thickness of a slipping zone w (see eqns [71]).

process zone can be significant even under conditions of small-scale yielding; in particular, for thin faults, the instantaneous temperature increase within the process zone is predicted to be a factor of σ_s/σ_d greater than the temperature increase on the rest of the slipping interface (Fialko, 2004a). Second, in the presence of a continued dynamic weakening (e.g., Abercrombie and Rice, 2005), the frictional heating is expected to progressively decay behind the rupture front, further suppressing the excess temperature. In the context of thermal weakening, such coupling between the coseismic heating and dynamic friction may be conducive to self-sustained slip pulses (Noda et al., 2009; Perrin et al., 1995; Zheng and Rice, 1998) and to a transition to a pulse-like behavior for ruptures that initially propagate in a cracklike mode.

Numerical simulations that allow for full coupling between the shear stress, shear heating, pore fluid pressure, and elasticity indicate a rich variety of slip behaviors and important controls exerted by the effective thickness of the fault slip zone w. For example, Garagash (2012) showed that in the presence of thermal pressurization, spontaneous fault slip can occur in the form of slip pulses, with the rupture speed V_r inversely correlated with the thickness of the slip zone. For the laboratory values of hydraulic and thermal properties of fault gouge, the model predicts earthquake-like behavior for slip zones having thickness of the order of centimeters (or thinner), an aseismic episodic slip with characteristics similar to those of the episodic slow slip events observed in subduction zones worldwide (Obara, 2002; Rogers and Dragert, 2003), assuming width of the slip zone of the order of 1 m (Garagash, 2012).

If thermal weakening mechanisms that operate at relatively small increases in the average fault temperature, such as thermal pressurization and flash melting (e.g., Andrews, 2002; Goldsby and Tullis, 2011; Lachenbruch, 1980; Lee and Delaney, 1987; Rice, 2006), do not give rise to substantial reductions in the dynamic friction, a continued dissipation and heating due to a localized slip will result in macroscopic melting on the slip interface and a transition from the asperity-contact friction to viscous rheology. High transient stresses associated with shearing of thin viscous films of melt have been considered as one of possible mechanisms of thermally induced strengthening (Fialko, 2004a; Koizumi et al., 2004; Tsutsumi and Shimamoto, 1997). However, recent laboratory, field, and theoretical studies suggest that transient viscous braking may not be an efficient arresting mechanism, especially in the lower part of the brittle layer (Brown and Fialko, 2012; Di Toro et al., 2006; Fialko and Khazan, 2005). If so, earthquakes that produced macroscopic melting and pseudotachylites (e.g., Davidson et al., 2003; Sibson, 1975; Swanson, 1992; Wenk et al., 2000) must have been accompanied by nearly complete stress drops. More generally, one may argue that highly localized slip zones are an indicator of nearly complete stress drops for sizeable earthquakes, regardless of whether macroscopic melting took place. This stems from the fact that melting of a narrow slip interface could have been prevented only if the dynamic friction were already low (Fialko, 2004a). Some theoretical arguments and experimental data (Brown and Fialko, 2012; Di Toro et al., 2006; Fialko and Khazan, 2005; Mitchell et al., 2013) also raise a question about the mechanism of rupture arrest below the brittle–ductile transition. A currently prevailing view is that the bottom of the seismogenic layer represents a rheological transition between the velocity-weakening friction and velocity-strengthening friction (e.g., Marone, 1998; Scholz, 1998). While this transition may well be caused by the temperature-controlled onset of stable creep (although recent laboratory data revealed a possibility of stick–slip behavior in granite at temperatures well above 300 °C, at least under dry conditions; see, e.g., Mitchell et al., 2013), it is unclear whether the velocity strengthening is relevant for the arrest of seismic ruptures that propagate into the ductile substrate from the brittle upper crust. For highly localized ruptures, the onset of melting or some other thermally activated mechanism may occur immediately behind the rupture front (aided by high ambient temperature and stress concentration), potentially overwhelming the effect of rate-and-state friction. In this case, the rupture arrest may require either delocalization of seismic slip near the rupture front (thereby limiting the temperature rise and increasing the effective fracture energy) or low deviatoric stress below the brittle–ductile transition. Discriminating between these possibilities may provide important insights into long-standing questions about the effective mechanical thickness and strength of tectonically active crust and lithosphere (England and Molnar, 1997; Jackson, 2002; Lamb, 2002; Takeuchi and Fialko, 2012).

4.03.9 Conclusions

Shear tractions on a slipping fault may be controlled by a multitude of physical processes that include yielding and breakdown around the propagating rupture front, rate- and slip-dependent friction, acoustic fluidization, dynamic reduction in the fault-normal stress due to dissimilar rigidities of the fault walls, thermally activated increases in the pore fluid pressure, flash heating of the asperities, and macroscopic melting. A traditional distinction between the fracture energy consumed near the rupture front and frictional losses on a well-slipped portion of a fault may be ill-defined if the slipping interface undergoes a substantial yet gradual weakening behind the rupture front. Both theoretical models and experimental data indicate that a continuous degradation of the dynamic fault strength is an expected consequence of a highly localized slip. In this case, earthquake ruptures violate a basic assumption of the LEFM, namely, that the zone of the material breakdown is small compared to the rupture size. The assumption of small-scale yielding is also violated if the extent of the off-fault damage is not negligible compared to the characteristic length of a slipping region. Field observations of kilometer-wide damage zones around major crustal faults (Ambraseys, 1970; Chen and Freymueller, 2002; Fialko, 2004b; Fialko et al., 2002) are evidence that the off-fault yielding may constitute a significant fraction of the earthquake energy balance (Wilson et al., 2004).

Both the dynamically induced variations in shear stress along the slip interface and distributed inelastic deformation off the primary slip surface affect the distribution of slip and slip rate, dynamic stress drop, and radiated seismic energy. Inferences from seismic data about constitutive properties of the fault zone material, such as the peak breakdown stress and critical slip-weakening distance, may be biased if the

aforementioned phenomena are not taken into account. In particular, simplified rupture models that neglect off-fault damage likely overestimate the critical slip-weakening distance D_c. Similarly, models that neglect the physics of high-speed slip may attribute the dynamic (e.g., thermally induced) weakening to large *in situ* values of D_c interpreted as an intrinsic rock fracture property. Constitutive relationships governing evolution of the dynamic strength at seismic slip velocities are difficult to formalize, as temperature, pore fluid pressure, and shear tractions on the fault interface are sensitive to a number of parameters such as the effective thickness of the slip zone, fluid saturation, dynamic damage and changes in the fault wall permeability, and ambient stress. Many of these parameters are also poorly constrained by the available data. Nonetheless, further insights into the dynamics of earthquakes may require fully coupled elastodynamic–thermodynamic models of seismic rupture. Development of such models is a significant challenge for future work.

Acknowledgments

I thank Jim Rice, Allan Rubin, and Hiroo Kanamori for their insightful reviews. Jim Rice kindly provided results of his calculations of off-fault damage (Figure 5). This work was supported by NSF through grants EAR-0338061, EAR-0450035, and EAR-0944336.

References

Abercrombie R (1995) Earthquake source scaling relationships from −1 to 5 M_L using seismograms recorded at 2.5-km depth. *Journal of Geophysical Research* 100: 24015–24036.

Abercrombie R and Rice JR (2005) Can observations of earthquake scaling constrain slip weakening? *Geophysical Journal International* 162: 406–424.

Aki K (1967) Scaling law of seismic spectrum. *Journal of Geophysical Research* 72: 1217–1231.

Aki K and Richards PG (1980) *Quantitative Seismology: Theory and Methods, I and II*, p. 577. San Francisco, CA: W. H. Freeman,.

Ambraseys N (1970) Some characteristic features of the North Anatolian fault zone. *Tectonophysics* 9: 143–165.

Ampuero J and Rubin AM (2008) Earthquake nucleation on rate and state faults – Aging and slip laws. *Journal of Geophysical Research* 113: B01302.

Andrews DJ (1976a) Rupture propagation with finite stress in antiplane strain. *Journal of Geophysical Research* 81: 3575–3582.

Andrews DJ (1976b) Rupture velocity of plane strain shear cracks. *Journal of Geophysical Research* 81: 5679–5687.

Andrews DJ (2002) A fault constitutive relation accounting for thermal pressurization of pore fluid. *Journal of Geophysical Research* 107: ESE 15–1–ESE 15–8. http://dx.doi.org/10.1029/2002JB001942.

Andrews DJ (2005) Rupture dynamics with energy loss outside the slip zone. *Journal of Geophysical Research* 110: B01307. http://dx.doi.org/10.1029/2004JB003191.

Aochi H and Madariaga R (2003) The 1999 Izmit, Turkey, earthquake: Nonplanar fault structure, dynamic rupture process, and strong ground motion. *Bulletin of the Seismological Society of America* 93: 1249–1266.

Archuleta RJ (1984) A faulting model for the 1979 Imperial Valley earthquake. *Journal of Geophysical Research* 89: 4559–4585.

Atkinson BK (1987) Introduction to fracture mechanics. In: Atkinson BK (ed.) *Fracture Mechanics of Rock*, pp. 1–23. London: Academic Press, chapter 1.

Barenblatt GI (1959) The formation of equilibrium cracks during brittle fracture. General ideas and hypotheses. Axially-symmetric cracks. *Journal of Applied Mathematics and Mechanics* 23: 622–636 (in Russian).

Barton N (1971) A relationship between joint roughness and joint shear strength. In: *Proceedings of International Symposium on Rock Mechanics, Nancy, Pap. I-8*.

Bazant ZP and Planas J (1998) *Fracture and Size Effect in Concrete and Other Quasibrittle Materials*. Boca Raton, FL: CRC Press.

Ben-Zion Y (2003) Appendix 2, key formulas in earthquake seismology. In: Lee W, Kanamori H, Jennings P, and Kisslinger C (eds.) *International Handbook of Earthquake and Engineering Seismology*, Vol. 81B, pp. 1857–1875. Amsterdam: Academic Press.

Ben-Zion Y and Rice JR (1997) Dynamic simulations of slip on a smooth fault in an elastic solid. *Journal of Geophysical Research* 102: 17771–17784.

Ben-Zion Y and Shi Z (2005) Dynamic rupture on a material interface with spontaneous generation of plastic strain in the bulk. *Earth and Planetary Science Letters* 236: 486–496.

Beroza GC and Mikumo T (1996) Short slip duration in dynamic rupture in the presence of heterogeneous fault properties. *Journal of Geophysical Research* 101: 22449–22460.

Beroza GC and Spudich P (1988) Linearized inversion for fault rupture. *Journal of Geophysical Research* 93: 6275–6296.

Bilby BA and Eshelby JD (1968) Dislocations and the theory of fracture. In: Liebowitz H (ed.). *Fracture: An Advanced Treatise*, vol. 1, pp. 99–182. New York: Academic Press.

Blanpied M, Tullis T, and Weeks J (1995) Frictional slip of granite at hydrothermal conditions. *Journal of Geophysical Research* 100: 13045–13064.

Bouchon M and Vallee M (2003) Observation of long supershear rupture during the magnitude 8.1 Kunlunshan earthquake. *Science* 301: 824–826.

Broberg K (1978) Transient sliding motion. *Geophysical Journal of the Royal Astronomical Society* 52: 397–432.

Broberg K (1987) On crack paths. *Engineering Fracture Mechanics* 28: 663–679.

Broberg K (1989) The near-tip field at high crack velocities. *International Journal of Fracture* 39: 1–13.

Broberg K (1999) *Cracks and Fracture*. London: Academic Press.

Brodsky EE and Kanamori H (2001) Elastohydrodynamic lubrication of faults. *Journal of Geophysical Research* 106: 16357–16374.

Brown KM and Fialko Y (2012) "Melt welt" mechanism of extreme weakening of gabbro at seismic slip rates. *Nature* 488: 638–641.

Brune JN, Henyey T, and Roy R (1969) Heat flow, stress, and rate of slip along San Andreas fault, California. *Journal of Geophysical Research* 74: 3821–4009.

Brune J and Thatcher W (2002) Strength and energetics of active fault zones. In: Lee W, Kanamori H, Jennings P, and Kisslinger C (eds.) *International Handbook of Earthquake and Engineering Seismology*, Vol. 81B, pp. 569–588. Amsterdam: Academic Press.

Burridge R (1973) Admissible speeds for plane-strain self-similar shear cracks with friction but lacking cohesion. *Geophysical Journal of the Royal Astronomical Society* 35: 439–455.

Burridge R and Halliday G (1971) Dynamic shear cracks with friction as models for shallow focus earthquakes. *Geophysical Journal of the Royal Astronomical Society* 25: 261–283.

Byerlee J (1978) Friction of rock. *Pure and Applied Geophysics* 116: 615–626.

Cardwell R, Chinn D, Moore G, and Turcotte D (1978) Frictional heating on a fault zone with finite thickness. *Geophysical Journal of the Royal Astronomical Society* 52: 525–530.

Chen Q and Freymueller J (2002) Geodetic evidence for a near-fault compliant zone along the San Andreas fault in the San Francisco Bay area. *Bulletin of the Seismological Society of America* 92: 656–671.

Chen YT and Knopoff L (1986) Static shear crack with a zone of slip-weakening. *Geophysical Journal of the Royal Astronomical Society* 87: 1005–1024.

Cherepanov GP (1968) Cracks in solids. *International Journal of Solids and Structures* 4: 811–831.

Cherepanov GP (1979) *Mechanics of Brittle Fracture*. Berlin, Heidelberg: Springer Verlag.

Chester JS, Chester FM, and Kronenberg AK (2005) Fracture surface energy of the Punchbowl fault, San Andreas system. *Nature* 437: 133–136.

Cowie PA and Scholz CH (1992) Physical explanation for the displacement–length relationship of faults using a post-yield fracture mechanics model. *Journal of Structural Geology* 14: 1133–1148.

Dahlen FA (1977) The balance of energy in earthquake faulting. *Geophysical Journal of the Royal Astronomical Society* 48: 239–261.

Dahlen FA and Tromp J (1998) *Theoretical Global Seismology*, p. 944. Princeton, NJ: Princeton University Press.

Davidson C, Davis K, Bailey C, Tape C, Singleton J, and Singer B (2003) Age, origin, and significance of brittle faulting and pseudotachylyte along the coast shear zone, prince rupert, British Columbia. *Geology* 31: 43–46.

Day SM (1982) Three-dimensional simulation of spontaneous rupture: The effect of nonuniform prestress. *Bulletin of the Seismological Society of America* 72: 1881–1902.

Delouis B, Giardini D, Lundgren P, and Salichon J (2002) Joint inversion of InSAR, GPS, teleseismic, and strong-motion data for the spatial and temporal distribution of earthquake slip: Application to the 1999 Izmit mainshock. *Bulletin of the Seismological Society of America* 92: 278–299.

Di Toro G, Goldsby DL, and Tullis TE (2004) Friction falls towards zero in quartz rock as slip velocity approaches seismic rates. *Nature* 427: 436–439.

Di Toro G, Hirose T, Nielsen S, Pennacchioni G, and Shimamoto T (2006) Natural and experimental evidence of melt lubrication of faults during earthquakes. *Science* 311: 647–649.

Dieterich JH (1979) Modeling of rock friction 1. Experimental results and constitutive equations. *Journal of Geophysical Research* 84: 2161–2168.

Dieterich JH (1992) Earthquake nucleation on faults with rate- and state-dependent strength. *Tectonophysics* 211: 115–134.

Dugdale DS (1960) Yielding of steel sheets containing slits. *Journal of the Mechanics and Physics of Solids* 8: 100–115.

Dunham EM and Archuleta RJ (2004) Evidence for a supershear transient during the 2002 Denali fault earthquake. *Bulletin of the Seismological Society of America* 94: S256–S268.

England PC and Molnar P (1997) Active deformation of Asia: From kinematics to dynamics. *Science* 278: 647–650.

Eshelby JD (1956) The continuum theory of lattice defects. In: Seitz F and Turnbull D (eds.) *Progress in Solid State Physics*, pp. 79–144. New York: Academic Press.

Fialko Y (2004a) Temperature fields generated by the elastodynamic propagation of shear cracks in the Earth. *Journal of Geophysical Research* 109: B01303. http://dx.doi.org/10.1029/2003JB002497.

Fialko Y (2004b) Probing the mechanical properties of seismically active crust with space geodesy: Study of the co-seismic deformation due to the 1992 M_w7.3 Landers (southern California) earthquake. *Journal of Geophysical Research* 109: B03307. http://dx.doi.org/10.1029/2003JB002756.

Fialko Y and Khazan Y (2005) Fusion by earthquake fault friction: Stick or slip? *Journal of Geophysical Research* 110: B12407. http://dx.doi.org/10.1029/2005JB003869.

Fialko Y, Rivera L, and Kanamori H (2005) Estimate of differential stress in the upper crust from variations in topography and strike along the San Andreas fault. *Geophysical Journal International* 160: 527–532.

Fialko YA and Rubin AM (1997) Numerical simulation of high pressure rock tensile fracture experiments: Evidence of an increase in fracture energy with pressure? *Journal of Geophysical Research* 102: 5231–5242.

Fialko Y, Sandwell D, Agnew D, Simons M, Shearer P, and Minster B (2002) Deformation on nearby faults induced by the 1999 Hector Mine earthquake. *Science* 297: 1858–1862.

Freund LB (1979) The mechanics of dynamic shear-crack propagation. *Journal of Geophysical Research* 84: 2199–2209.

Freund LB (1998) *Dynamic Fracture Mechanics, p. 563*. New York: Cambridge University Press.

Gakhov FD (1966) *Boundary Value Problems*. London: Pergamon Press.

Garagash D (2012) Seismic and aseismic slip pulses driven by thermal pressurization of pore fluid. *Journal of Geophysical Research* 117: B04314.

Goldsby D and Tullis T (2002) Low frictional strength of quartz rocks at subseismic slip rates. *Geophysical Research Letters* 29: 25–1–25–4. http://dx.doi.org/10.1029/2002GL015240.

Goldsby D and Tullis T (2011) Flash heating leads to low frictional strength of crustal rocks at earthquake slip rates. *Science* 334: 216–218.

Griffith AA (1920) The phenomena of rupture and flow in solids. *Philosophical Transactions of the Royal Society of London* A221: 163–197.

Gu J-C, Rice JR, Ruina AL, and Tse ST (1984) Slip motion and stability of a single degree of freedom elastic system with rate and state dependent friction. *Journal of the Mechanics and Physics of Solids* 32: 167–196.

Gutenberg B and Richter C (1949) *Seismicity of the Earth and Associated Phenomena*. Princeton: Princeton University Press.

Hartzell SH and Heaton T (1983) Inversion of strong ground motion and teleseismic waveform data for the fault rupture history of the 1979 Imperial Valley, California, earthquake. *Bulletin of the Seismological Society of America* 73: 1553–1583.

Hashida T, Oghikubo H, Takahashi H, and Shoji T (1993) Numerical-simulation with experimental-verification of the fracture-behavior in granite under confining pressures based on the tension-softening model. *International Journal of Fracture* 59: 227–244.

Heaton T (1990) Evidence for and implications of self-healing pulses of slip in earthquake rupture. *Physics of the Earth and Planetary Interiors* 64: 1–20.

Hirose T and Shimamoto T (2003) Fractal dimension of molten surfaces as a possible parameter to infer the slip-weakening distance of faults from natural pseudotachylytes. *Journal of Structural Geology* 25: 1569–1574.

Hirose T and Shimamoto T (2005) Growth of molten zone as a mechanism of slip weakening of simulated faults in gabbro during frictional melting. *Journal of Geophysical Research* 110: B05202. http://dx.doi.org/10.1029/2004JB003207.

Husseini MI (1977) Energy balance for formation along a fault. *Geophysical Journal of the Royal Astronomical Society* 49: 699–714.

Ida Y (1972) Cohesive force across the tip of a longitudinal-shear crack and Griffith's specific surface energy. *Journal of Geophysical Research* 77: 3796–3805.

Inglis CE (1913) Stresses in a plate due to the presence of cracks and sharp corners. *Transactions of the Royal Institution of Naval Architects* 55: 219–241.

Ingraffea AR and Schmidt RA (1978) Experimental verification of a fracture mechanics model for tensile strength prediction of Indiana limestone. In: Kim Y-S (ed.) *Proceedings of the 19th US Symposium on Rock Mechanics*, pp. 247–253. Reno: University of Nevada.

Irwin GR (1957) Relation of stresses near a crack to the crack extension force. In: *Proceedings of 9th International Congress on Applied Mechanics, Vol. VIII*, pp. 245–251.

Jackson J (2002) Strength of the continental lithosphere: Time to abandon the jelly sandwich? *GSA Today* 12: 4–9.

Jeffreys H (1942) On the mechanics of faulting. *Geological Magazine* 79: 291–295.

Kame N and Yamashita T (1999) Simulation of the spontaneous growth of a dynamic crack without constraints on the crack tip path. *Geophysical Journal International* 139: 345–358.

Kanamori H (2004) The diversity of the physics of earthquakes. *Proceedings of the Japan Academy, Series B* 80: 297–316.

Kanamori H and Anderson DL (1975) Theoretical basis of some empirical relations in seismology. *Bulletin of the Seismological Society of America* 65: 1073–1095.

Kanamori H and Brodsky EE (2004) The physics of earthquakes. *Reports on Progress in Physics* 67: 1429–1496.

Kanamori H and Heaton TH (2000) Microscopic and macroscopic physics of earthquakes. In: Rundle J, Turcotte DL, and Klein W (eds.) *Physics of Earthquakes*. *Geophysical Monograph*, vol. 106, pp. 117–136. Washington, DC: AGU.

Khazan YM and Fialko YA (1995) Fracture criteria at the tip of fluid driven cracks in the Earth. *Geophysical Research Letters* 22: 2541–2544.

Khazan YM and Fialko Y (2001) Tensile and shear cracks in the Dugdale–Barenblatt model. *Geofizicheskii Zhurnal* 23: 13–30 (in Russian).

Koizumi YK, Otsuki TA, and Nagahama H (2004) Frictional melting can terminate seismic slips: Experimental results of stick-slips. *Geophysical Research Letters* 31. http://dx.doi.org/10.1029/2004GL020, 642, Art. No. L21,605.

Kolosoff GV (1909) *On an Application of Complex Function Theory to a Plane Problem of the Mathematical Theory of Elasticity*. Yuriev.

Kostrov BV (1970) Tectonic earthquake focal theory. *Izvestiya Akademii Nauk SSSR, Earth Physics* 4: 84–101.

Kostrov BV (1974) Seismic moment and energy of earthquakes and seismic flow of rock. *Izvestiya, Academy of Sciences, USSR, Physics of the Solid Earth (Engl. Translation)* 1: 23–40.

Kostrov BV and Das S (1988) *Principles of Earthquake Source Dynamics*, p. 286. Cambridge: Cambridge University Press.

Lachenbruch AH (1980) Frictional heating, fluid pressure, and the resistance to fault motion. *Journal of Geophysical Research* 85: 6097–6112.

Lamb S (2002) Is it all in the crust? *Nature* 420: 130–131.

Landau LD and Lifshitz EM (1986) *Theory of Elasticity*, p. 187. Oxford: Pergamon Press.

Lapusta N and Rice JR (2004) Earthquake sequences on rate and state faults with strong dynamic weakening. *EOS, Transactions, American Geophysical Union, Supplement* 84: S51B02.

Lawn B (1993) *Fracture of Brittle Solids*, 2nd edn, p. 378. Cambridge: Cambridge University Press.

Lee T-C and Delaney P (1987) Frictional heating and pore pressure rise due to fault slip. *Geophysical Journal of the Royal Astronomical Society* 88: 569–591.

Leonov MY and Panasyuk VV (1959) The development of very small cracks in a solid. *Zhurnal Prikladnoi Mekhaniki i Tekhnicheskoi Fiziki (Applied Mechanics)* 5: 391–401 (in Ukrainian).

Li VC (1987) Mechanics of shear rupture applied to earthquake zones. In: Atkinson B (ed.) *Fracture Mechanics of Rock*, pp. 351–428. London: Academic Press.

Li Y, Aki K, Vidale J, and Alvarez M (1998) A delineation of the Nojima fault ruptured in the M7.2 Kobe, Japan, earthquake of 1995 using fault zone trapped waves. *Journal of Geophysical Research* 103(B4): 7247–7263.

Lockner DA, Byerlee J, Kuksenko V, Ponomarev A, and Sidorin A (1992) Observations of quasistatic fault growth from acoustic emissions. In: Evans B and Wong T (eds.) *Fault Mechanics and Transport Properties of Rocks*, pp. 3–31. San Diego, CA: Academic.

Madariaga R (1976) Dynamics of an expanding circular fault. *Bulletin of the Seismological Society of America* 66: 636–666.

Maddock R (1986) Frictional melting in landslide-generated frictionites (hyalomylonites) and fault-generated pseudotachylytes – Discussion. *Tectonophysics* 128: 151–153.

Mair K and Marone C (2000) Shear heating in granular layers. *Pure and Applied Geophysics* 157: 1847–1866.

Malvern LE (1969) *Introduction to the Mechanics of a Continuum Medium*, p. 713. Englewood Cliffs, NJ: Prentice-Hall.
Manighetti I, King G, Gaudemer Y, Scholz C, and Doubre C (2001) Slip accumulation and lateral propagation of active normal faults in Afar. *Journal of Geophysical Research* 106: 13667–13696.
Marone C (1998) Laboratory-derived friction laws and their application to seismic faulting. *Annual Review of Earth and Planetary Sciences* 26: 643–696.
Mase CW and Smith L (1987) Effects of frictional heating on the thermal, hydrologic, and mechanical response of a fault. *Journal of Geophysical Research* 92: 6249–6272.
McKenzie D and Brune JN (1972) Melting on fault planes during large earthquakes. *Geophysical Journal of the Royal Astronomical Society* 29: 65–78.
Melin S (1986) When does a crack grow under mode II conditions. *International Journal of Fracture* 30: 103–114.
Mitchell E, Fialko Y, and Brown KM (2013) Temperature dependence of frictional healing of Westerly granite: Experimental observations and numerical simulations. *Geochemistry, Geophysics, Geosystems* 14: 567–582.
Molinari A, Estrin Y, and Mercier S (1999) Dependence of the coefficient of friction on the sliding conditions in the high velocity range. *Journal of Tribology* 121: 35–41.
Muskhelishvili NI (1953) *Some Basic Problems of Mathematical Theory of Elasticity*. Groningen, Holland: Noordhoff.
Noda H, Dunham E, and Rice JR (2009) Earthquake ruptures with thermal weakening and the operation of major faults at low overall stress levels. *Journal of Geophysical Research* 114: B07,302.
Obara K (2002) Nonvolcanic deep tremor associated with subduction in southwest Japan. *Science* 296: 1679.
Oglesby DD and Archuleta RJ (2001) Fault geometry and the dynamics of the 1999 Chi-Chi (Taiwan) earthquake. *Bulletin of the Seismological Society of America* 91: 1099–1111.
Ohnaka M (2003) A constitutive scaling law and a unified comprehension for frictional slip failure, shear fracture of intact rock, and earthquake rupture. *Journal of Geophysical Research* 108: B2080. http://dx.doi.org/10.1029/2000JB000,123.
Okada Y (1985) Surface deformation due to shear and tensile faults in a half-space. *Bulletin of the Seismological Society of America* 75: 1135–1154.
Olsen K, Madariaga R, and Archuleta R (1997) Three-dimensional dynamic simulation of the 1992 Landers earthquake. *Science* 278: 834–838.
Palmer AC and Rice JR (1973) The growth of slip surfaces in the progressive failure of over-consolidated clay. *Proceedings of the Royal Society of London, Series A* 332: 527–548.
Perrin G, Rice JR, and Zheng G (1995) Self-healing slip pulse on a frictional surface. *Journal of the Mechanics and Physics of Solids* 43: 1461–1495.
Peyrat S, Olsen K, and Madariaga R (2001) Dynamic modeling of the 1992 Landers earthquake. *Journal of Geophysical Research* 106: 26467–26482.
Poliakov AB, Dmowska R, and Rice JR (2002) Dynamic shear rupture interactions with fault bends and off-fault secondary faulting. *Journal of Geophysical Research* 107. http://dx.doi.org/10.1029/2001JB000,572.
Power WL and Tullis TE (1995) A review of the fractal character of natural fault surfaces with implications for friction and the evolution of fault zones. In: Barton CC and LaPointe PR (eds.) *Fractals in the Earth Sciences*, pp. 89–105. New York: Plenum.
Reches Z and Lockner DA (2010) Fault weakening and earthquake instability by powder lubrication. *Nature* 467. http://dx.doi.org/10.1038/nature09348 (452-U102).
Reid H (1910) The mechanism of the earthquake. In: Comm RSI (ed.). *The California Earthquake of April 19, 1906*, vol. 2 Washington, DC: Carnegie Institution.
Rice JR (1968a) Mathematical analysis in the mechanics of fracture. In: Liebowitz H (ed.). *Fracture: An Advanced Treatise*, vol. 2, pp. 191–311. New York: Academic Press.
Rice JR (1968b) A path-independent integral and the approximate analysis of strain concentration by notches and cracks. *Journal of Applied Mechanics* 35: 379–386.
Rice JR (1979) The mechanisms of quasistatic crack growth. In: Kelly RE (ed.) *Proceedings of 8th US National Congress on Applied Mechanics, UCLA, June 1978*, pp. 191–216. North Hollywood, CA: Western Periodicals.
Rice JR (1980) The mechanics of earthquake rupture. In: Dziewonski AM and Boschi E (eds.) *Physics of the Earth's Interior*, pp. 555–649. Amsterdam: North-Holland.
Rice JR (1993) Spatio-temporal complexity of slip on a fault. *Journal of Geophysical Research* 98: 9885–9907.
Rice JR (2006) Heating and weakening of faults during earthquake slip. *Journal of Geophysical Research* 111: B05,311.
Rice JR, Sammis CG, and Parsons R (2005) Off-fault secondary failure induced by a dynamic slip pulse. *Bulletin of the Seismological Society of America* 95: 109–134.
Richards P (1976) Dynamic motions near an earthquake fault – 3-dimensional solution. *Bulletin of the Seismological Society of America* 66: 1–32.
Rivera LA and Kanamori H (2005) Representations of the radiated energy in earthquakes. *Geophysical Journal International* 162: 148–155.
Rogers G and Dragert H (2003) Episodic tremor and slip on the Cascadia subduction zone: The chatter of silent slip. *Science* 300: 1942.
Rubin AM (1993) Tensile fracture of rock at high confining pressure: Implications for dike propagation. *Journal of Geophysical Research* 98: 15919–15935.
Rubin AM and Ampuero J-P (2005) Earthquake nucleation on (aging) rate and state faults. *Journal of Geophysical Research* 110: B11,312. http://dx.doi.org/10.1029/2005JB003,686.
Rudnicki JW (1980) Fracture mechanics applied to the Earth's crust. *Annual Review of Earth and Planetary Sciences* 8: 489–525.
Rudnicki JW and Freund LB (1981) On energy radiation from seismic sources. *Bulletin of the Seismological Society of America* 71: 583–595.
Ruina A (1983) Slip instability and state variable friction laws. *Journal of Geophysical Research* 88: 10359–10370.
Samudrala O, Huang Y, and Rosakis AJ (2002) Subsonic and intersonic shear rupture of weak planes with a velocity weakening cohesive zone. *Journal of Geophysical Research* 107: B2170. http://dx.doi.org/10.1029/2001JB000,460.
Savage J (1998) Displacement field for an edge dislocation in a layered half-space. *Journal of Geophysical Research* 103: 2439–2446.
Scholz CH (1998) Earthquakes and friction laws. *Nature* 391: 37–42.
Scholz CH (2002) *The Mechanics of Earthquakes and Faulting*, 2nd edn, p. 496. New York, NY: Cambridge University Press.
Segall P (1991) Fault mechanics. *Reviews of Geophysics* 29: 864–876.
Segall P and Rice JR (1995) Dilatancy, compaction, and slip instability of a fluid infiltrated fault. *Journal of Geophysical Research* 100: 22155–22171.
Sibson RH (1973) Interaction between temperature and pore-fluid pressure during earthquake faulting – A mechanism for partial or total stress relief. *Nature* 243: 66–68.
Sibson RH (1975) Generation of pseudotachylite by ancient seismic faulting. *Geophysical Journal of the Royal Astronomical Society* 43: 775–794.
Sibson R (1980) Power dissipation and stress levels on faults in the upper crust. *Journal of Geophysical Research* 85: 6239–6247.
Simonov IV (1983) Behavior of solutions of dynamic problems in the neighborhood of the edge of a cut moving at transonic speed in the elastic medium. *Mechanics of Solids (Mekhanika Tverdogo Tela)* 18: 100–106.
Spray JG (2005) Evidence for melt lubrication during large earthquakes. *Geophysical Research Letters* 32. http://dx.doi.org/10.1029/2004GL022,293, Art. No. L07,301.
Steketee JA (1958) Some geophysical applications of the elasticity theory of dislocations. *Canadian Journal of Physics* 36: 1168–1198.
Swanson MP (1992) Fault structure, wear mechanisms and rupture processes in pseudotachylyte generation. *Tectonophysics* 204: 223–242.
Takeuchi C and Fialko Y (2012) Dynamic models of interseismic deformation and stress transfer from plate motion to continental transform faults. *Journal of Geophysical Research* 117: B05,403. http://dx.doi.org/10.1029/2011JB009,056.
Timoshenko S and Goodier JN (1970) *Theory of Elasticity*, 3rd edn. New York: McGraw-Hill, 398 pp.
Tinti E, Spudich P, and Cocco M (2005) Earthquake fracture energy inferred from kinematic rupture models on extended faults. *Journal of Geophysical Research* 110: B12,303. http://dx.doi.org/10.1029/2005JB003,644.
Tsutsumi A and Shimamoto T (1997) High-velocity frictional properties of gabbro. *Geophysical Research Letters* 24: 699–702.
Vvedenskaya AV (1959) Determination of displacement fields for earthquakes by means of the dislocation theory. *Izvestiya Akademii nauk SSSR; Seriya geofizicheskaya* 3: 277–284 (in Russian).
Wald DJ and Heaton TH (1994) Spatial and temporal distribution of slip for the 1992 Landers, California, earthquake. *Bulletin of the Seismological Society of America* 84: 668–691.
Wenk H-R, Johnson L, and Ratschbacher L (2000) Pseudotachylites in the Eastern Peninsular Ranges of California. *Tectonophysics* 321: 253–277.
Westergaard H (1939) Bearing pressures and cracks. *Journal of Applied Mechanics, Transactions of the ASME* 6: A49–A53.
Willis JR (1967) A comparison of the fracture criteria of Griffith and Barenblatt. *Journal of the Mechanics and Physics of Solids* 15: 151–162.
Wilson B, Dewers T, Reches Z, and Brune J (2004) Particle size and energetics of gouge from earthquake rupture zones. *Nature* 434: 749–752.
Xia K, Rosakis AJ, and Kanamori H (2004) Laboratory earthquakes: The sub-Rayleigh-to- supershear rupture transition. *Science* 303: 1859–1861.
Zheng G and Rice J (1998) Conditions under which velocity-weakening friction allows a self-healing versus a cracklike mode of rupture. *Bulletin of the Seismological Society of America* 88: 1466–1483.

This page intentionally left blank

4.04 Applications of Rate- and State-Dependent Friction to Models of Fault-Slip and Earthquake Occurrence

JH Dieterich, University of California, Riverside, CA, USA

© 2007 Elsevier B.V. All rights reserved.

This chapter is reproduced from the previous edition, Volume 4, pp. 107–129, © 2007 Elsevier B.V.

4.04.1	**Fault-Slip Phenomena**	93
4.04.2	**Rate- and State-Dependent Friction**	93
4.04.3	**Models of Sliding Phenomena**	95
4.04.3.1	Stability of Slip	95
4.04.3.2	Depth Dependence of Constitutive Parameters	96
4.04.3.3	Minimum Fault Dimension for Unstable Slip	96
4.04.3.4	Creep Events and Silent Earthquakes	97
4.04.3.5	Earthquake Afterslip	99
4.04.4	**Earthquake Nucleation**	99
4.04.4.1	Numerical Simulations of Spontaneous Earthquake Nucleation	99
4.04.4.2	Nucleation Solutions for Single Degree-of-Freedom Slider	101
4.04.5	**Seismicity Rate Models**	101
4.04.5.1	Earthquake Rate Formulation	101
4.04.5.2	Aftershocks	102
4.04.5.2.1	Aftershock-rate equations	102
4.04.5.2.2	Variability of aftershock decay rates	103
4.04.5.3	Foreshocks	103
4.04.5.4	Earthquake Triggering	103
4.04.6	**Stress Changes Estimated from Earthquake Rates**	106
4.04.7	**Conclusions and Future Directions**	107
References		108

4.04.1 Fault-Slip Phenomena

Slip of active faults takes a variety of forms including continuous creep, episodic creep events or slow earthquakes, earthquake slip, decelerating postseismic creep, and inference localized accelerating slip during earthquake nucleation. Any comprehensive theory of fault-slip and earthquake occurrence must be based on a physical understanding of these diverse sliding phenomena. Sliding behaviors analogous to those of active faults are also seen in laboratory fault-slip experiments, where sliding processes and constitutive properties may be studied directly. The rate- and state-dependent constitutive formulation describes characteristic dependencies of friction on slip, slip rate, slip history, and normal stress history observed in laboratory studies, and it provides a unified framework for predictive modeling of the various sliding phenomena observed for faults in nature and in the laboratory. A formulation for earthquake rates, which is derived using rate- and state-dependent friction, forms the basis of an integrated approach to quantitatively model the effects of stress changes on earthquake rates. Foreshocks, aftershocks, and triggering phenomena can be modeled; and the earthquake-rate formulation has been used in an inverse mode to infer stress changes from changes of seismicity rates.

4.04.2 Rate- and State-Dependent Friction

The sliding resistance of rocks approximately conforms to the Coulomb friction model, wherein the coefficient of friction μ is a constant. The nominal coefficient of friction of rocks, like other unlubricated materials, is generally about 0.6–0.7 (Byerlee, 1978; Stesky, 1978). However, laboratory studies have established that in detail, the coefficient friction is not constant – there are subtle but characteristic dependencies on slip, sliding rate, contact time, and normal stress history (Beeler et al., 1994; Blanpied and Tullis, 1986; Dieterich, 1979, 1981; Linker and Dieterich, 1992; Marone, 1998; Ruina, 1983; Tullis, 1988; Tullis and Weeks, 1986). The rate- and state-dependent constitutive formulation quantifies these frictional dependencies, providing a framework to unify observations of dynamic/static friction, displacement weakening at the onset of macroscopic slip, time-dependent healing, slip history dependence, and slip speed dependence (Dieterich and Kilgore, 1996a; Ruina, 1983). Although developed expressly for fault slip, rate–state constitutive properties appear to be characteristic of most unlubricated materials – this has led to its use in diverse applications including the analysis of friction in microelectromechanical system (MEMS) devices, servo-control design for precise positioning systems, failure processes in composite materials, and even the design of clamps in window frames.

Several closely related rate- and state-dependent formulations have been used to study sliding phenomena and earthquake processes (Dieterich, 1979, 1981; Rice, 1983; Ruina, 1983). The Ruina (1983) simplification of the Dieterich (1981) formulation for sliding resistance is widely used and may be written as

$$\tau = \sigma\left[\mu_0 + a\ln\left(\frac{\dot{\delta}}{\dot{\delta}^*}\right) + b\ln\left(\frac{\theta}{\theta^*}\right)\right] \qquad [1]$$

 http://dx.doi.org/10.1016/B978-0-444-53802-4.00075-0

where τ and σ are shear stress and effective normal stress (compression-positive), respectively; μ_0, a, and b are experimentally determined constants; $\dot{\delta}$ is the sliding speed; θ is a state variable (discussed later in the text) that evolves with slip- and normal stress history; and $\dot{\delta}^*$ and θ^* are normalizing constants. The constant θ^* in eqn [1] is often replaced by $\theta^* = D_c/\dot{\delta}^*$ where D_c is a characteristic slip parameter described in the succeeding text. The nominal coefficient of friction, μ_0, is defined at a reference slip rate and state ($\dot{\delta} = \dot{\delta}^*, \theta = \theta^*$ and generally has values of 0.6–0.7). For silicates at room temperature, a and b have roughly similar values in the range 0.005–0.015 (Dieterich, 1981; Kilgore et al., 1993; Linker and Dieterich, 1992; Marone, 1998; Marone and Kilgore, 1993; Ruina, 1983; Tullis and Weeks, 1986).

As $\dot{\delta}$ or θ approach zero, eqn [1] yields unacceptably small (or negative) values of sliding resistance. To limit the minimum value of μ, some studies (Dieterich, 1986, 1987; Okubo, 1989; Okubo and Dieterich, 1986; Shibazaki and Iio, 2003) use a modified form of eqn [1] that adds constants to the logarithmic terms:

$$\tau = \sigma\left[\mu_0 + a\ln\left(\frac{\dot{\delta}}{\dot{\delta}^*} + 1\right) + b\ln\left(\frac{\theta}{\theta^*} + 1\right)\right] \qquad [2]$$

Here, μ_0 takes the value of the minimum friction. However, at present, very few experimental data exist that let us assign limiting values to $\dot{\delta}^*$ or θ^*. In granitic rocks, Okubo and Dieterich (1986) analyzed recordings of slip speed and frictional strength at the tips of propagating frictional instabilities in granite and detected a lower limit for the state effect with an average value $\theta^* = 0.03$ s. However, controlled sliding experiments with granite to speeds as low as 10^{-4} μm s^{-1} failed to detect a lower limit to the rate effect (Kilgore et al., 1993).

The state variable θ has the dimensions of time and is interpreted to be a measure of the age of the population load-supporting contacts, which evolve with contact time, slip, and changes of normal stress (Dieterich, 1979, 1981; Dieterich and Kilgore, 1994). The connection between the age of contacts and frictional strength appears to originate from localized creep, which occurs under the extreme stress conditions that exist at microcontacts. Microscopic imaging of contacts in transparent materials (including minerals, glass, and plastics) reveals that the actual area of contact increases with the logarithm of contact time (Dieterich and Kilgore, 1994, 1996b). The contact creep observations are closely related to the well-known microindentation creep effect (Dieterich and Conrad, 1984; Dieterich and Kilgore, 1994, 1996b; Evans and Goetz, 1979; Goldsby et al., 2004; Scholz and Engelder, 1976; Walker and Demer, 1964; Westbrook and Jorgensen, 1965).

Different empirical evolution laws for the state variable θ have been proposed (Blanpied and Tullis, 1986; Linker and Dieterich, 1992; Perrin et al., 1995; Ruina, 1983). Common features of the laws are evolution of state over a characteristic slip distance D_c and steady-state dependence of state on slip rate whereby contact population age is inversely proportional to slip speed:

$$\theta_{ss} = D_c/\dot{\delta} \qquad [3]$$

The evolution law that is used in the following discussion incorporates time-dependent evolution of stationary contacts, which accounts for observed strengthening with the logarithm of time during stationary contact (Dieterich, 1972), displacement evolution proposed by Ruina (1983), and normal stress dependence of Linker and Dieterich (1992). It is

$$d\theta = dt - \frac{\theta}{D_c}d\delta - \frac{\alpha\theta}{b\sigma}d\sigma \qquad [4]$$

where α is a constant $0 \le \alpha \le \mu$. Measurements using bare surfaces of Westerly granite give α in the range 0.25–0.50 (Linker and Dieterich, 1992). Characteristic slip distance D_c increases with surface roughness, gouge particle size, and gouge thickness and generally varies from 2 to 100 μm (Dieterich, 1981; Dieterich and Kilgore, 1994; Marone, 1998; Tullis and Weeks, 1986). Under conditions of constant normal stress, the evolution laws have the property that θ decreases with time at stresses above the steady-state friction, which can, under circumstances described later in the text, lead to frictional weakening and unstable (earthquake) slip. Conversely, θ increases at stresses below the steady-state friction, which leads to strengthening of the frictional resistance. This strengthening at lower stresses provides a mechanism for recovery of fault strength following unstable slip events.

At constant normal stress, the steady-state dependence of eqn [3] combined with eqn [1] gives the following relation for the coefficient of friction at steady state:

$$\mu_{ss} = \mu_0 + (a-b)\ln\left(\frac{\dot{\delta}}{\dot{\delta}^*}\right) \qquad [5]$$

Equation [5] employs the definition $\theta^* = D_c/\dot{\delta}^*$. With this substitution, μ_0 now represents the steady-state friction at slip, the reference speed $\dot{\delta}^*$. Note that if $a > b$, then friction increases with increasing slip speed (rate strengthening), and if $a < b$, then friction decreases with increasing slip speed (rate weakening).

The interplay between the rate-dependent and state-dependent effects and the evolution of state over the distance D_c are revealed in the characteristic response of the coefficient of friction to rapid steplike transitions from one sliding speed to another (**Figure 1**). In **Figure 1**, the top curve gives the idealized response from the rate- and state-dependent constitutive formulation and the lower curves give the observed response for a variety of materials. Step changes of sliding speed, under conditions of constant normal stress, result in immediate jumps in frictional resistance followed by displacement-dependent decay and stabilization at a new steady-state sliding friction. In rocks, this characteristic rate- and state-dependent frictional behavior has been documented for bare surfaces and simulated fault surfaces separated by a layer of gouge under a wide range of conditions accessible in laboratory experiments. These include wet and dry conditions, the range of temperatures and pressures characteristic of crustal earthquakes, and sliding rates from mm year^{-1} to mm s^{-1} (Blanpied and Tullis, 1986; Blanpied et al., 1991; Dieterich, 1979, 1981; Kilgore et al., 1993; Linker and Dieterich, 1992; Marone, 1998; Marone and Kilgore, 1993; Ruina, 1983; Tullis, 1988; Tullis and Weeks, 1986; Weeks and Tullis, 1985).

Rate- and state-dependent effects described here are largely based on laboratory studies conducted at slow slip speeds. Hence, they are clearly appropriate for modeling fault creep processes and earthquake nucleation. Experimental data for friction during stick-slip events with small displacements (Okubo and Dieterich, 1986) and plate impacts with very high slip speeds (tens of meters per second) (Prakash, 1998) indicate that rate–state behaviors also operate at seismic slip speeds. Models of the earthquake cycle that employ rate–state effects

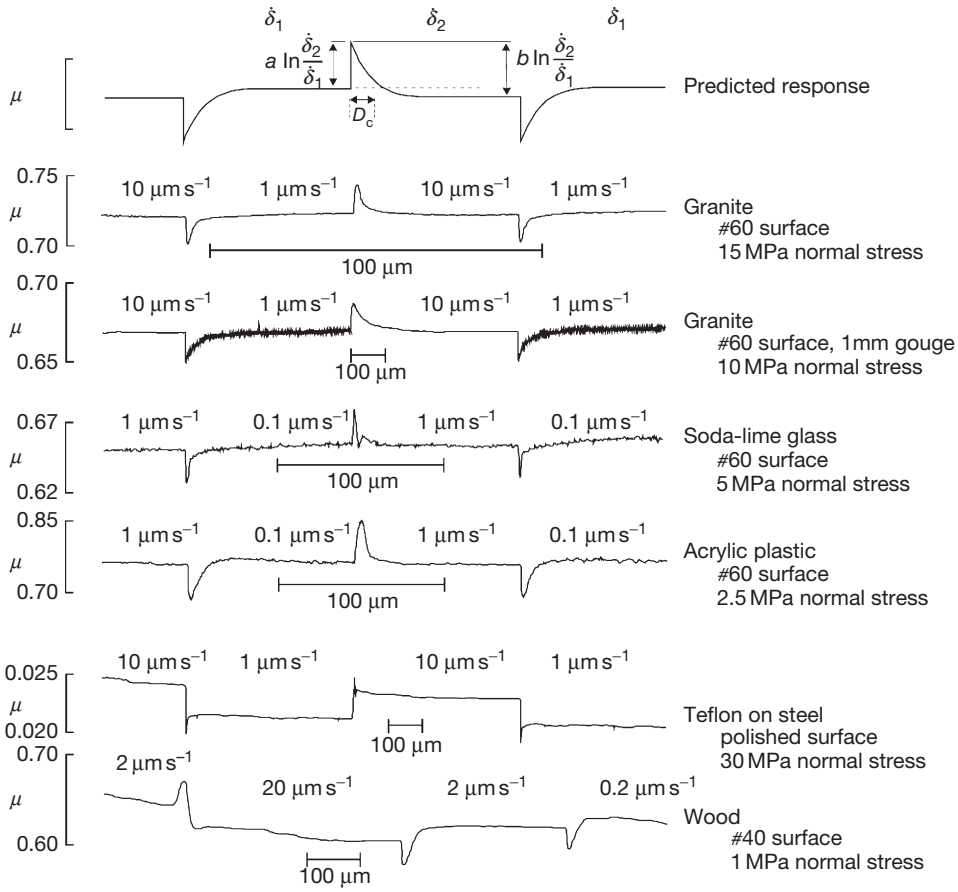


Figure 1 Characteristic rate–state frictional response to step changes of slip speed. Remarkably, a similar behavior is seen for a wide variety of materials. Top curve indicates the predicted rate–state response. D_c is observed to depend on surface roughness and characteristics of gouge separating the surfaces (Dieterich and Kilgore, 1996b).

during earthquake slip (Kato and Hirasawa, 1997; Lapusta and Rice, 2003; Lapusta et al., 2000; Okubo, 1989; Stuart, 1988; Tse and Rice, 1986) yield results that are consistent with a broad scope of observations. However, the conditions and processes that control sliding resistance during the large displacements and high slip speeds of large earthquakes remain uncertain. Processes that may significantly alter, or replace, the rate- and state-dependent effects include thermal interactions due to shear heating and pore-fluid effects (Rice, 2006; Segall and Rice, 1995; Sleep, 1995a,b, 1997; Sleep and Blanpied, 1992). These and other processes, together with fault constitutive properties for slip in large earthquakes, are the subject of **Chapter 4.05**.

4.04.3 Models of Sliding Phenomena

4.04.3.1 Stability of Slip

Sliding instabilities are a common, but generally unwelcome, phenomenon in a variety of mechanical systems that vary in scale from microns in MEMS devices to a thousand kilometers or more with earthquake faults. The conditions required for unstable slip with rate- and state-dependent friction have been extensively analyzed for a single degree of freedom system consisting of a slider connected to a loading point through a spring of stiffness K (Blanpied and Tullis, 1986; Dieterich and Linker, 1992; Gu et al., 1984; Rice and Ruina, 1983; Ruina, 1983; Weeks and Tullis, 1985). Here, stiffness is defined as the change of stress at the slider divided by change in spring elongation. For sliders at constant normal stress, those studies establish that unstable slip occurs only when the system stiffness is less than the critical value

$$K_c = \frac{\xi\sigma}{D_c} \qquad [6]$$

where ξ is a parameter that depends on loading conditions and constitutive parameters. For perturbations from slow steady-state slip at constant normal stress, $\xi = b - a$ (Rice and Ruina, 1983; Ruina, 1983). This result is independent of the evolution law used. The terms K_c, σ, and D_c are positive – hence, an added requirement for unstable slip is that ξ must also be positive (i.e., $b > a$). This is readily understood from eqn [5], which shows that steady-state frictional resistance decreases with increasing slip rate if $b > a$. Conversely, the strengthening that occurs with increasing slip speed if $b < a$ will act to prevent acceleration of slip leading to instability. Laboratory studies confirm the association of stick slip with rate weakening and additionally verify predictions of the critical stiffness for instability, including the dependence on σ and D_c. Stable slip always occurs if $b > a$, or if $K > K_c$ independent of the relative values of a and b.

In many circumstances, including slip of nonplanar faults and nonvertical faults that interact with the free surface, both normal and shear components of stress couple to slip. The Ruina stability analysis has been generalized to include cases where both shear and normal stresses are coupled to slip through a spring inclined at an angle to the sliding surface (Dieterich and Linker, 1992). This leads to two stability criteria. The first is a generalized form of the Ruina and Rice results, in which the term ξ in the critical stiffness equation is

$$\xi = \frac{b-a}{1+(\mu_{ss}-\alpha)\tan\phi}, \quad (\text{for } \phi > -\cot^{-1}\mu_{ss}) \qquad [7]$$

where α is the normal stress parameter in the evolution eqn [4] and ϕ is the angle between the spring and the sliding surface (taken as positive if slip increases the compressive stress). For the case $\phi = 0$, normal stress remains constant during slip. It is seen from this result that conditions in which normal stress decreases with increasing slip (negative spring angles) are less stable than situations in which normal stress increases with slip; that is, negative spring angles result in larger ξ than positive spring angles and therefore require a stiffer spring to stabilize the slip. However, the requirement of $b > a$ for instability remains valid (for $\phi > -\cot^{-1}(\mu_{ss})$). The second stability condition arises for large negative spring angles. If $\phi < -\cot^{-1}(\mu_{ss})$, then either slider lockup or instability occurs. The second condition is independent of the system stiffness or specific constitutive properties (i.e., b may be less than a).

4.04.3.2 Depth Dependence of Constitutive Parameters

Laboratory data on steady-state friction at different temperatures (Blanpied et al., 1991, 1995; Stesky, 1978) indicate that rocks with rate-weakening friction at low temperatures undergo a transition to rate-strengthening friction above temperatures of about 300–350 °C (**Figure 2(a)**). Tse and Rice (1986) noted that the condition for slip instability, $b > a$ given in eqn [6], and the temperature dependence of steady-state friction are consistent with the observed depth ranges of earthquakes along plate boundaries; that is, the predicted depth of the isotherm for the transition from unstable slip to fault creep, based on thermal gradients at plate boundaries, corresponds to the observed maximum depth of earthquakes. These observations indicate a model for the depth dependence of a and b (**Figure 2(b)**) that is consistent with the depth range of seismicity, fault stability criteria, and predictions of constitutive properties based on crustal temperatures. The temperature-dependent model for constitutive parameters has been widely adopted for modeling the earthquake cycle and creep processes on strike-slip faults and in subduction zones (Kato and Hirasawa, 1997; Lapusta and Rice, 2003; Lapusta et al., 2000; Stuart, 1988; Tse and Rice, 1986). Shallow portions of faults with rate strengthening are sometimes invoked to represent poorly consolidated sediments in the shallow portions of subducting slabs and crustal strike-slip faults. This association is based on observations of rate strengthening at low normal stresses in poorly consolidated layers of gouge (Marone, 1998; Marone et al., 1990). Such a shallow zone is consistent with observations of shallow fault creep, which tends to correlate with poorly consolidated sediments.

4.04.3.3 Minimum Fault Dimension for Unstable Slip

An approximate but useful representation of effective stiffness for slip of a fault patch, obtained from the elasticity solutions for slip of a crack embedded in an elastic media, takes the form

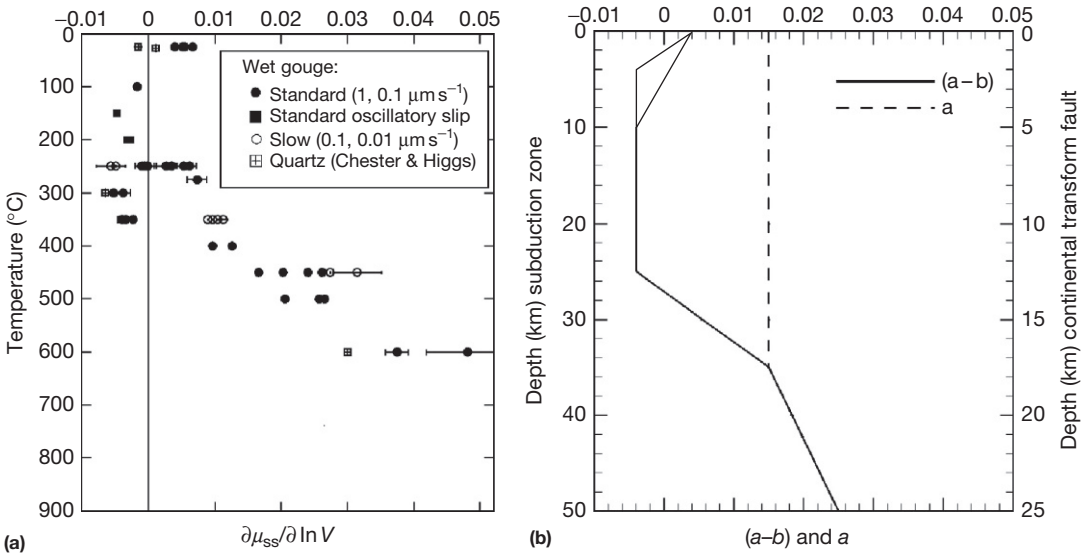


Figure 2 Depth distribution of friction parameters. (a) Experimental data on steady-state friction under hydrothermal conditions from Blanpied et al. (1995). (b) Depth distribution of (a–b) and a based on thermal structure of continental transform faults and subduction zones. The shallow rate-strengthening zone may be of variable thickness depending on the degree of consolidation of sediments. Subduction of cold sediments may produce zones of rate weakening at greater depths than generally found along strike-slip faults. Adapted from Liu Y and Rice JR (2005) Aseismic slip transients emerge spontaneously in three-dimensional rate and state modeling of subduction earthquake sequences. *Journal of Geophysical Research* 110: B08307, http://dx.doi.org/10.1029/2004JB003424; Lapusta N and Rice JR (2003) Nucleation and early seismic propagation of small and large events in a crustal earthquake model. *Journal of Geophysical Research* 108(B4): 2205, http://dx.doi.org/10.1029/2001JB000793.

$$K = \frac{\Delta\tau}{\delta} = \frac{G\eta}{L} \quad [8]$$

where L is the length (or diameter) of the fault patch, G is the shear modulus (Poisson's ratio taken to be 0.25), and η is a factor with value near 1 that depends upon the geometry of the slip patch and slip or stress conditions on the patch. Table 1 gives values of η for some simple patch geometries. Combining eqn [8] with the critical stiffness relation for instability (eqn [6]) gives

$$L_c = \frac{G\eta D_c}{\xi\sigma} \quad [9]$$

where L_c is the patch length corresponding to K_c (Dieterich, 1986, 1992). Slip of fault patches with dimensions less than L_c will have an effective stiffness in excess of K_c and sliding will be stable.

Assuming that the dimensions of an earthquake rupture must be equal to, or greater than, the size of the patch that nucleates the instability, L_c defines a lower bound for the dimensions of the earthquake source. Equation [9] has been approximately verified by experiments with a large-scale testing apparatus that accommodates samples with a 2 m fault – a size that is sufficient to permit confined sliding instabilities that do not reach to the ends of the fault (Dieterich and Kilgore, 1996a).

A primary uncertainty in applying eqn [9] to faults in nature relates to poor constraints on the parameter D_c, which continues to be an open question. In the laboratory, D_c increases with surface roughness, gouge particle size, and possibly gouge thickness. Compared to natural faults, laboratory experiments employ exceptionally smooth faults with thin layers of fine gouge. It seems plausible, therefore, that natural faults may have much larger values of D_c than laboratory faults, which could substantially increase the estimates of the minimum dimensions of earthquakes that originate as frictional instabilities. However, the existence of very small earthquakes with source dimensions of 5 m or less (Abercrombie, 1995) requires small nucleation sources and values of D_c, which may be similar to those measured in laboratory experiments. For example, using an estimate of D_c at the lower end of laboratory values (2 μm) gives $L_c \sim 5$ cm ($G = 18\,000$ MPa, $\xi = 0.004$, $\eta = 1$, with effective $\sigma = 100$ MPa). Increasing D_c to the upper range of laboratory values (100 μm) gives $L_c \sim 2$ m.

4.04.3.4 Creep Events and Silent Earthquakes

Aseismic creep of active faults is observed in a variety of tectonic settings where it takes the form of either continuous creep or creep events. Continuous creep at more or less steady rates is readily understood in terms of the stability relations discussed in Section 4.04.3.3, which show that steady creep will occur when $a > b$. However, the origins of creep events are more problematic.

Creep events were first observed with surface creepmeters along portions of the creeping section of the San Andreas Fault in central California (Nason and Weertman, 1973; Wesson, 1988). The events there are periodic, have durations that vary from an hour to a few days, and have displacements of up to a few tens of millimeters per slip event. Analyses of strain signals associated with these events indicate that the sources are shallow, extending to maximum depths of a few kilometers (Gladwin et al., 1994; Langbein, 1981; McHugh and Johnston, 1976).

Worldwide improvements in monitoring capabilities have revealed widespread occurrence of deep creep events. The events span a very large range of sizes (equivalent to moment magnitudes up to $M \sim 7$) and have durations varying from rapid short-term slip that is transitional with earthquake slip to very slow events that continue more than a year. The terms slow earthquake and silent earthquake are sometimes used to refer to short-term and long-term events, respectively. Creep events of various durations have been reported in widespread locations and geologic settings including Japan (Fujii, 1993; Hirose et al., 1999; Kawasaki, 2004; Linde and Sacks, 2002; Ozawa et al., 2002; Sacks et al., 1981; Sagiya, 2004), California (Kanamori and Hauksson, 1992; Linde and Sacks, 2002), Alaska (Freymueller et al., 2002), Cascadia (Dragert et al., 2001; Miller et al., 2002; Rogers and Dragert, 2003), Italy (Crescentini et al., 1999), Mexico (Kostoglodov et al., 2003; Lowry et al., 2001), Hawaii at Kīlauea volcano (Brooks et al., 2006; Cervelli et al., 2002), and oceanic transform faults (Beroza and Jordan, 1990). Repeating slip events have been reported for Cascadia and Hawaii. The Hawaiian events appear to be highly periodic (Brooks et al., 2006; Miller et al., 2002).

The continuous spectrum of event durations from minutes to years suggests a common origin for creep events. In very general terms, the characteristics of creep events resemble those of an earthquake cycle in that both have long intervals in which the fault is essentially locked (at least to the resolution of deformation observations), followed by an interval of rapid slip. From a mechanical perspective, creep events require a weakening mechanism to enable rapidly accelerating slip and some type of damping mechanism to quench the acceleration of slip before it reaches seismic slip speeds. In the case of earthquakes, acceleration of slip is made possible by slip-rate weakening ($a < b$), and slip speed is limited by inertial damping from the generation of seismic waves.

Table 1 Value of parameter η for different patch geometries[a]

η	L	*Patch*	*Reference*
$7\pi/12$	Diameter	Circular patch	Eshelby (1957)
4/3	Length	Plane strain, constant shear stress	Star (1928)
$8/3\pi$	Length	Plane strain, constant slip	Chinnery (1969)
1	Length (vertical)	Antiplane strain, constant shear stress	Knopoff (1958)
$2/\pi$	Length (vertical)	Antiplane strain, constant slip	Chinnerry (1969)

[a]Assuming Poisson's ratio $\nu = 0.25$, slip δ is measured at the center of the patch.

Oscillatory creep is observed in laboratory experiments and in spring-slider simulations near the stable–unstable slip transition of eqn [6]. Near this boundary, small decreases of effective normal stress can move a system from a conditionally stable state to an unstable state. Kodaira et al. (2004) had identified a region of high pore-fluid pressure, a subduction zone where silent slip occurs in the Tōkai (Japan) region. They proposed that high pore-fluid pressures extend the region of conditionally stable slip to greater depths, which promotes creep events over large regions. However, the necessary conditions for oscillations and the amplitudes of the fluctuations of slip speeds are very restricted in single degree-of-freedom sliders.

Recent 3-D simulations of earthquakes and creep in subduction zones were found to spontaneously generate larger slip pulses in the region of transition from seismic to aseismic slip when mild heterogeneities are introduced (Liu and Rice, 2005). The simulated creep events have some characteristics of observed creep events. However, in their current form, these simulations do not produce the highly periodic events with relatively short periods seen in some areas, nor do they address the small shallow periodic surface creep events seen in the creeping section of the San Andreas Fault system. It is not yet clear if this approach can reproduce the spectrum of slip rates and slip amplitudes evident in natural systems.

Shibazaki and Iio (2003) proposed a model in which creep events arise spontaneously from the modified form of the rate–state formulation given by eqn [2]. In particular, the use of a lower cutoff for θ together with $b > a$ produces a transition from steady-state rate weakening to rate strengthening at $\theta^* = \theta_{ss} = D_c/\dot{\delta}$. At slow speeds, rate weakening prevails, which provides the necessary mechanism for acceleration of slip, and at higher speeds, the fault experiences rate strengthening, which provides the necessary mechanism to halt the acceleration before seismic slip speeds are reached. Characteristics of creep events with this mechanism are readily illustrated with a spring-slider model using a rate–state law with a single cutoff in the state term

$$\tau = \sigma\left[\mu_0 + a\ln\left(\frac{\dot{\delta}}{\dot{\delta}^*}\right) + b\ln\left(\frac{\theta}{\theta^*} + 1\right)\right] \qquad [10]$$

Figure 3 illustrates steady-state friction from eqn [10] and the oscillatory response of a slider with these properties. This simple version of the Shibazaki and Iio (2003) model

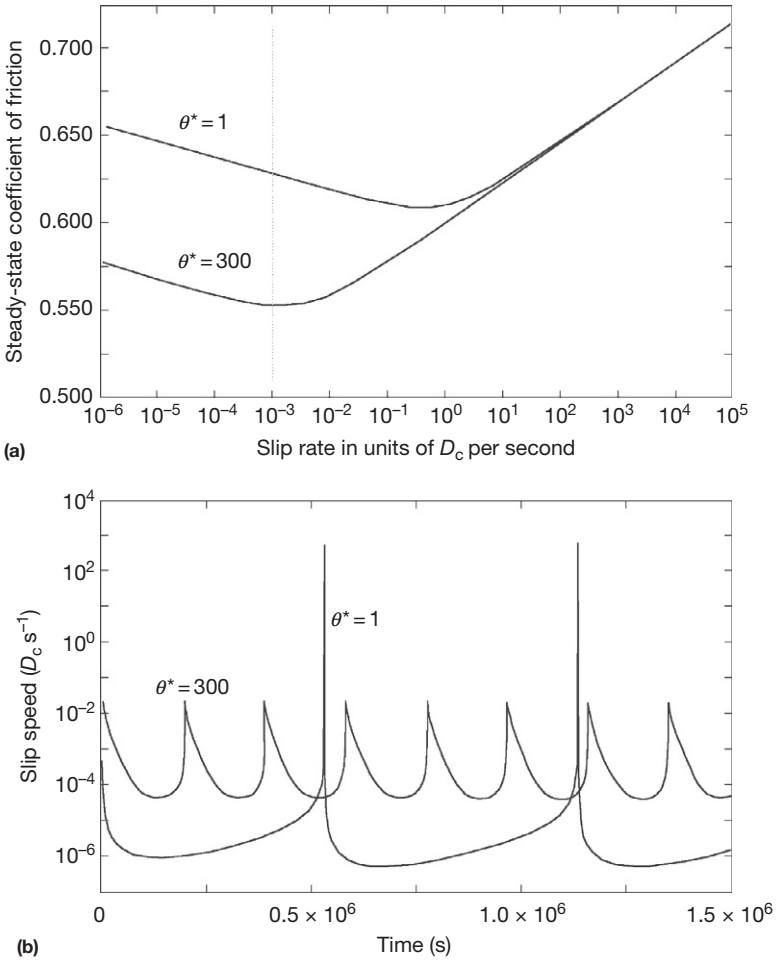


Figure 3 Spring-slider simulation of creep events. (a) Steady-state friction as given by eqn [10], which produces a minimum in friction around $\dot{\delta} = D_c/\theta^*$. Dotted line is load point displacement rate. Units of θ^* are in seconds. (b) The minimum in friction produces oscillatory slip, which increases in amplitude and period as the minimum friction moves to increasing higher speeds.

produces highly periodic creep events. By changing θ^*, the duration and amplitude of the creep events are continuously adjustable from slow low-amplitude events, to rapid events, to fully seismic slip. However, at present, there are no direct experimental data to either justify or discount the use cutoffs at the values required to produce oscillations resembling those in nature. Most experiments to evaluate rate–state friction effects have used unaltered quartzofeldspathic materials that give relatively strong rate weakening at slow slip speeds and room temperature. There are few data for the highly altered gouges and sediments that are likely to occur either in zones of shallow creep or at high temperatures that characterize deep silent earthquakes.

4.04.3.5 Earthquake Afterslip

Earthquake afterslip is an episode of stable slip that characteristically increases with the logarithm of time following an earthquake. It was originally recognized through direct measurements of fault offset at the Earth's surface, and it generally occurs in regions where thick poorly consolidated sediments lie above seismogenic faults (Bucknam et al., 1978; Burford, 1976; Smith and Wyss, 1968). In addition, deep afterslip is inferred from postseismic strain observations (Bürgmann et al., 2002; Hsu et al., 2002; Hutton et al., 2002; Miyazaki et al., 2004; Pollitz et al., 1998). Both shallow and deep afterslips are well represented by models that incorporate rate- and state-dependent friction. However, surface deformation observations may be insufficient to differentiate deep afterslip from bulk viscoelastic flow in some circumstances.

In general terms, significant afterslip occurs whenever earthquake slip increases the stresses acting on an adjacent fault segment that has rate-strengthening characteristics ($a>b$). Models of the seismic cycle on strike-slip faults and subduction zones that employ a deep zone with rate-strengthening friction (**Figure 2**) spontaneously generate afterslip following earthquake slip events (Kato and Hirasawa, 1997; Lapusta and Rice, 2003; Lapusta et al., 2000; Stuart, 1988; Tse and Rice, 1986). Those studies show that earthquake slip is strongly impeded when it encounters regions where $b>a$, which results in large stress increases and accelerated postearthquake creep rates. Creep rates decay as the afterslip relaxes the stresses.

Marone et al. (1991) and Wennerberg and Sharp (1997) modeled surface afterslip as a single degree-of-freedom slider. In the case of shallow afterslip, Marone et al. (1991) pointed out that poorly consolidated fault gouges have rate-strengthening characteristics. Their models fit afterslip observations with considerable fidelity and incorporate normal stress and stiffness parameters scaled to the depth of afterslip in natural fault systems. Because the seismogenic portion of the fault is essentially locked following an earthquake, the slider represents the afterslip zone, and it is assigned rate-strengthening friction. The spring stiffness scales by eqn [6], where L is the thickness of the afterslip zone. Earthquake slip distorts the spring and loads the slider, which responds by slipping at a faster rate. Both studies assume a quasi steady-state condition, which assumes slip is always at steady-state. This appears to be a reasonable approximation because slip speed changes slowly and slip is large (centimeters to meters) compared to laboratory values of D_c (2–100 μm). Earthquake slip (spring load point displacement) has a finite duration and causes some coseismic slip of the slider, which is readily calculated. This fixes the stress and initial afterslip speed. Marone et al. (1991) gave the following solution for afterslip with a spring and slider:

$$\delta = \frac{\sigma(a-b)}{K} \ln\left[\left(\frac{K\dot{\delta}_{CS}}{\sigma(a-b)}\right)t+1\right] + \dot{\delta}_0 t \qquad [11]$$

where t is the time from the end of the earthquake, $\dot{\delta}_{CS}$ is the thickness-averaged slip speed in the rate-strengthening area at $t=0$, and $\dot{\delta}_0$ is the long-term creep rate in the rate-strengthening area.

4.04.4 Earthquake Nucleation

4.04.4.1 Numerical Simulations of Spontaneous Earthquake Nucleation

The events and interactions that lead to the initiation of unstable fault slip are collectively defined as earthquake nucleation (Dieterich, 1987, 1992). The characterization of the earthquake nucleation process is of considerable interest because nucleation establishes the time and place of occurrence of earthquakes, which in turn determine their space–time patterns.

Dieterich (1992) and Rubin and Ampuero (2005) examined models with heterogeneous initial conditions, uniform loading, and uniform constitutive properties, and Lapusta and Rice (2003) modeled nucleation at the base of a seismogenic zone where there are strong stressing gradients and spatial variations of constitutive properties. Each of the studies obtains qualitatively similar results, though details vary as a consequence of differences in model configurations and initial conditions. **Figure 4** illustrates some characteristic features of the nucleation process. Nucleation arises spontaneously and

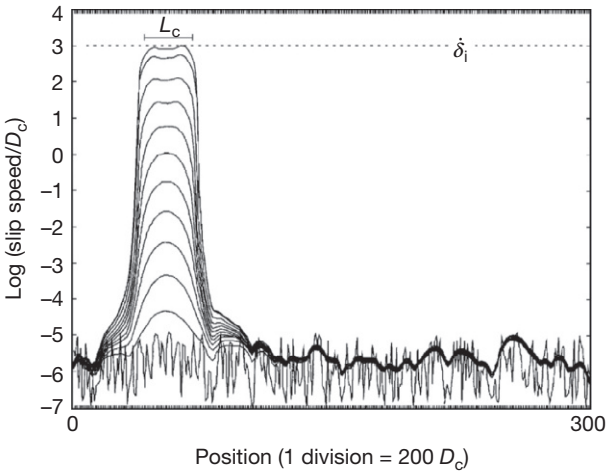


Figure 4 Model of spontaneous earthquake nucleation on a fault with rate- and state-dependent friction (Dieterich, 1992). The model was initialized with random normal stress and uniform shear stress. Growth of slip speeds favors wavelengths around L_c.

consists of an extended interval of stable accelerating slip, during which slip speed increases by $1/t$, where t is the time remaining to instability. As nucleation proceeds, it rapidly localizes to a portion of the fault and becomes increasingly insensitive to external stressing. Dieterich (1992) and Rubin and Ampuero (2005) found a region of model parameter space where accelerating slip remains localized to a zone of fixed length L_c, which scales by eqn [9] with $\xi \approx 0.4b$. In simulations where slip is forced to begin over a region $L > L_c$, the region of most active slip contracts to L_c as the time of instability approaches. In addition, there is a region of parameter space where the initial localization of accelerating slip is followed by expansion of the nucleation zone prior to unstable slip. This type of nucleation is characteristic of the Lapusta and Rice (2003) models. Dieterich (1992) and Rubin and Ampuero (2005) obtained different results on the extent of the parameter space where final expansion occurs, most likely as

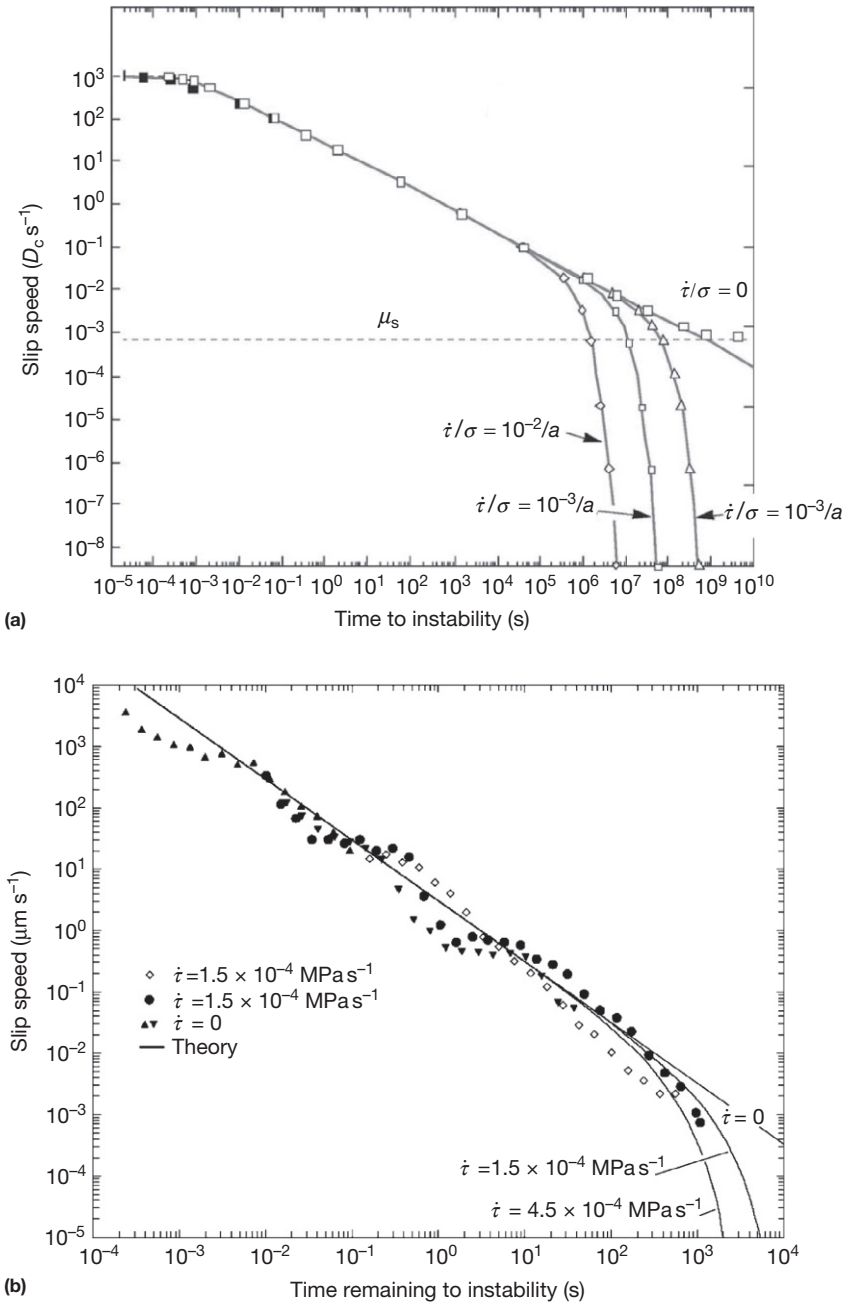


Figure 5 Time to nucleate versus slip speed. (a) Comparison of numerical calculations with 2-D boundary element model illustrated in **Figure 4** (data points), with analytic solutions for spring and slider (solid lines) as given by eqns [17] and [20] (Dieterich, 1992). (b) Comparison of experimental results for nucleation on 2 m laboratory fault (data points) with spring–slider solutions (Dieterich and Kilgore, 1996a). The triangle represents experiments in which external stressing was stopped when premonitory displacements were detected, demonstrating the self-driven nature of the nucleation process.

a consequence of differences in the way the computations are initiated.

4.04.4.2 Nucleation Solutions for Single Degree-of-Freedom Slider

Analytic solutions for a simplified model, which represents the nucleation source as a fault patch of fixed length, are given by Dieterich (1992). The localization process observed in the numerical models allows the use of eqn [8] to represent slip of the patch as a single degree-of-freedom slider with spring stiffness K. The model pertains to conditions that operate once nucleation is in progress $\left(\theta >> D_c/\dot{\delta}\right)$. Although several approximations are employed, the solutions are in good agreement with both detailed 2-D simulations and laboratory observations of the nucleation process on a 2 m fault (**Figure 5**). These solutions form the basis of the earthquake-rate formulation discussed in **Section 4.04.5** and are useful for understanding earthquake-triggering processes.

In the following, the Dieterich (1992) solutions have been generalized using the evolution eqn [4] for conditions in which normal stress is not constant. A general expression for acceleration of slip during nucleation is given by

$$d\omega = \frac{-1}{a\sigma}\left[b\,dt + \omega d\tau + \left(\frac{\tau}{\sigma} - \alpha\right)\omega d\sigma\right] \qquad [12]$$

where $\omega = 1/\dot{\delta}$ and b is a term containing model and constitutive parameters

$$b = -K + \frac{b\sigma}{D_c} \qquad [13]$$

Assuming the terms $a\sigma$ and $b\sigma$ are constant (small $\Delta\sigma/\sigma$) and taking $\mu' = \tau/\sigma - \alpha$ to be an effective constant coefficient of friction during nucleation, eqn [12] reduces to

$$d\omega = \frac{-1}{a\sigma}[b\,dt + \omega dS] \qquad [14]$$

where S is a modified Coulomb stress function

$$S = \tau - \mu'\sigma \qquad [15]$$

Some solutions of eqn [14] are listed in **Table 2**. Solutions in terms of shear stress, with constant σ, are identical to those in **Table 2** with the replacement ($\tau = S$). The solutions in **Table 2** require specification of slip speed as an initial condition. Hence, the response of a nucleation source may be followed for a complex stressing history made up of any number of stressing ramps and stress steps, by using the slip-rate solutions to track the initial conditions (slip speed) through each time/stress step.

4.04.5 Seismicity Rate Models

4.04.5.1 Earthquake Rate Formulation

From eqn [1], it is evident that changes of stress result in changes of slip speed. The nonlinearity of this effect, together with the dependence of nucleation time on slip speed (eqns [18] and [21]), indicate that earthquake rates will be sensitive to stress changes. Effects of stress changes on earthquake rates have been modeled by treating seismicity as a sequence of nucleation events in which the stressing history controls the timing of earthquakes (Dieterich, 1979, 1994). This conceptual approach has been implemented with numerical models that follow the evolution of single degree-of-freedom sliders to the point of instability, using perturbed and unperturbed stressing, to explicitly determine the effects of stress steps, earth tides, and seismic transients on the discrete nucleation sources (Belardinelli et al., 2003; Dieterich, 1979, 1987; Gomberg et al., 1997, 1998, 2000).

Following a similar conceptual approach, a general analytic formulation for earthquake rates has been derived using the solutions for earthquake nucleation in **Table 2** (Dieterich, 1994). This formulation avoids the necessity of following individual nucleation sources, and it is valid for stressing histories of arbitrary complexity. Earthquake rate R in some magnitude interval is given by

$$R = \frac{r}{\gamma\dot{\tau}_r} \qquad [22]$$

$$d\gamma = \frac{1}{a\sigma}\left[dt - \gamma d\tau - \gamma\left(\frac{\tau}{\sigma} - \alpha\right)d\sigma\right] \qquad [23]$$

where r is a constant steady-state background rate at the reference shear stressing rate $\dot{\tau}_r$ and γ is a variable that evolves with time and stress. Note that constitutive parameters b and D_c that appear in the nucleation solutions cancel out of the rate formulation. Assuming small changes of stress ($a\sigma$ and τ/σ constant), eqns [22] and [23] can be written using the modified Coulomb stress function of eqn [15] (Dieterich et al., 2000, 2003):

Table 2 Nucleation solutions[a] (eqn [14])

	$\dot{S} = const \neq 0$	$\dot{S} = 0$
Slip	$\delta = \frac{-a}{h}\ln\left\{\frac{\dot{\delta}_0 h\sigma}{\dot{S}}\left[1 - \exp\left(\frac{\dot{S}t}{a\sigma}\right)\right] + 1\right\}$ [16]	$\delta = \frac{-a}{h}\ln\left\{1 - \frac{\dot{\delta}_0 ht}{a}\right\}$ [19]
Slip speed	$\dot{\delta} = \left\{\left[\frac{1}{\dot{\delta}_0} + \frac{h\sigma}{\dot{S}}\right]\left[\exp\left(\frac{-\dot{S}t}{a\sigma}\right)\right] - \frac{h\sigma}{\dot{S}}\right\}^{-1}$ [17]	$\dot{\delta} = \left\{\frac{1}{\dot{\delta}_0} - \frac{-ht}{a}\right\}^{-1}$ [20]
Time to instability[b]	$t_{EQ} = \frac{a\sigma}{\dot{S}}\ln\left[\frac{\dot{S}}{h\sigma\dot{\delta}_0} + 1\right]$ [18]	$t_{EQ} = \frac{a}{h}\left[\frac{1}{\dot{\delta}_0}\right]$ [21]

[a]These solutions are subject to the requirement that evolution of state does not reach the steady condition, $\left(\theta_{ss} = D_c/\dot{\delta}\right)$, prior to the onset of seismic slip (see Dieterich, 1992, for discussion). Initial condition is slip speed $\dot{\delta}_0$ at $t = 0$.

[b]Because inertial effects are ignored in this simplified model, instability is defined as the time where $\dot{\delta}_0 \to \infty$. Time to instability t_{EQ} is the time remaining to instability given δ_0.

$$R = \frac{r}{\gamma \dot{S}_r} \quad [24]$$

$$d\gamma = \frac{1}{a\sigma}[dt - \gamma dS] \quad [25]$$

For positive shear stressing rates with $d\sigma = 0$, eqn [23] has the property that γ seeks the steady-state value, $\gamma_{ss} = 1/\dot{\tau}$ with the characteristic relaxation time of $t_a = a\sigma/\dot{\tau}$. Similarly in terms of Coulomb stress, $\gamma_{ss} = 1/\dot{S}$ and $t_a = a\sigma/\dot{S}$.

In forward models of earthquake rates, one must find γ as a function of time for a specified stressing history. Earthquake rates as a function of time are then directly obtained from eqn [22] or [24]. Solutions for $\gamma(t)$ as a function of time may be found through numerical evaluation of the evolution equations. Alternatively, any complex stressing history may be reduced to a series of small time increments with constant stressing rates and stress steps. This permits the use of analytic solutions of the evolution equations (**Table 3**) to follow the evolution of γ through the sequence of stress/time steps.

4.04.5.2 Aftershocks

4.04.5.2.1 Aftershock-rate equations

Dieterich (1986, 1992) proposed that aftershocks are caused by the steplike changes of stress that occur at the time of a mainshock, which perturbs earthquake nucleation times. Two general observations support this model. First, aftershocks cluster in those areas near a mainshock where stress changes favor fault slip in aftershocks (Harrts et al., 1995; King et al., 1994; Reasenberg and Simpson, 1992; Stein, 1999; Toda et al., 1998). Second, rapid increases of stress unrelated to an earthquake mainshock also stimulate bursts of seismicity that obey the $1/t$ Omori aftershock decay law. Examples of the latter are given by the aftershock-like earthquake sequences at Kīlauea volcano that begin at the time of magmatic intrusion events (Dvorak et al., 1986).

Solutions of the earthquake-rate equations as a function of time following a step of Coulomb stress are given in **Table 4** and **Figure 6**. These results are obtained through the stepwise application of the evolution equations in **Table 3**. Solutions for shear stressing at constant normal stress are identical to those in **Table 4**, with the substitution $\tau = S$.

Equation [33] can be written in a form that is identical the Omori aftershock decay law:

$$R = \frac{c_1}{c_2 + t} \quad [35]$$

where the empirical parameters in the Omori equation are expressed in terms of constitutive parameters and stressing conditions in this model:

Table 3 Solutions of evolution equations for γ[a]

	Equation [23]	*Equation [25]*[b]
Evolution with time, fixed stress	$\gamma = \gamma_0 + \frac{t}{a\sigma}$	$\gamma = \gamma_0 + \frac{t}{a\sigma}$ [26]
$\dot{\tau} = \text{const} \neq 0$ $\dot{S} = \text{const} \neq 0$	$\gamma = \left(\gamma_0 - \frac{1}{\dot{\tau}}\right)\exp\left(\frac{-t}{t_a}\right) + \frac{1}{\dot{\tau}}$ [27]	$\gamma = \left(\gamma_0 - \frac{1}{\dot{S}}\right)\exp\left(\frac{-t}{t_a}\right) + \frac{1}{\dot{S}}$ [29]
Stress step	$\gamma = \gamma_0\left(\frac{\sigma}{\sigma_0}\right)\exp\left(\frac{\tau_0}{a\sigma_0} - \frac{\tau}{a\sigma}\right)$ [28]	$\gamma = \gamma_0 \exp\left(\frac{S_0 - S}{a\sigma}\right)$ [30]
log t stressing[c] $S = S_0 + u\ln(wt+1)$	$\gamma = \gamma_0(wt+1)^{-m} + \frac{(wt+1) - (wt+1)^{-m}}{a\sigma\, w(m+1)}$,	where $m = \frac{u}{a\sigma}$ [31]

[a]Parameter values at the start of time/stress increment are indicated by the subscript values 0. At the start of a time increment, $t = 0$.
[b]Solutions in terms of τ are at constant normal stress σ.
[c]Identical solution for γ is obtained for change of τ by log of time with constant normal stress.

Table 4 Earthquake rates following a stress step[a]

	Solution in terms of S
$\dot{\tau}_1 \neq 0, \dot{S}_1 \neq 0$	$R = \frac{r\dot{S}_1/\dot{S}_0}{\left[\frac{\dot{S}_1}{\dot{S}_0}\exp(-\Delta S/(a\sigma)) - 1\right]\exp(-t/t_a) + 1}$, where $t_a = a\sigma/\dot{S}_1$ [32]
$\dot{\tau}_1 = 0, \dot{S}_1 = 0$	$R = \frac{r}{\exp(-\Delta S/a\sigma) + \frac{t\dot{S}_0}{a\sigma}}$ [33]
Log stressing[b]	$R = \frac{r}{(wt+1)^{-m}\exp\left(\frac{-\Delta S}{a\sigma}\right) + \frac{(wt+1) - (wt+1)^{-m}}{\dot{S}_0 a\sigma\, w(m+1)}}$, where $m = \frac{u}{a\sigma}$ [34]

[a]Parameter values at the start and end of time/stress increment are indicated by the subscript values 0 and 1, respectively. Time $t = 0$ at the time of the stress step ΔS.
[b]Stressing by log of time following stress step as given in **Table 3**, eqn [31].

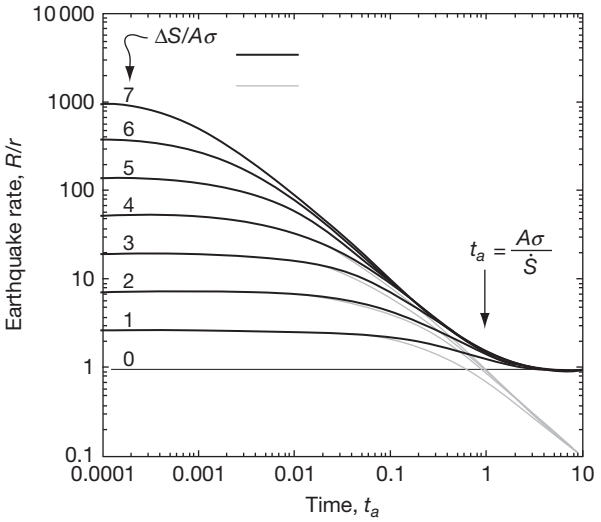


Figure 6 Plot of aftershock rate from eqn [32] (black lines), which has continuing stressing at $\dot{S}$ following a stress step, and from eqn [33] (gray lines), which has $\dot{S} = 0$ following the step.

$$c_1 = \frac{ra\sigma}{\dot{S}_r} \quad \text{and} \quad c_2 = \frac{ra\sigma}{\dot{S}_r}\exp\left(\frac{-\Delta S}{a\sigma}\right) \qquad [36]$$

Equation [31] (**Table 4** and **Figure 6**), which incorporates continued stressing after the stress step, also has the form of the Omori aftershock decay law at times $t < t_a$, and earthquake rates asymptotically return to a constant steady-state background at times $t > t_a$. Hence, t_a is the aftershock duration, which is inversely proportional to stressing rate. This inverse dependence on stressing rate is supported by observations of aftershock sequences (**Figure 7**), which have durations that span a range time from a few days to 78 years (Dieterich, 1994; Toda et al., 2002).

4.04.5.2.2 Variability of aftershock decay rates

Aftershocks decay at $1/t^p$, generally with $p \sim 1$. However, deviations from $p = 1$ appear to be reliable features of aftershock data. With the exception of eqn [34], which is discussed later in the text, the stress step equations of **Table 4** predict aftershock decay exactly at $p = 1$. However, predicted aftershock rates integrated over a region with heterogeneous stress changes generally give decay with $p \neq 1$. For example, $p \approx 0.8$ is obtained by integrating aftershock rates over a finite region that has stress variations characteristic of those adjacent to a shear crack (Dieterich, 1994). Also, slip of a nonplanar fault with random fractal roughness results in highly heterogeneous stress changes adjacent to the fault (**Figure 8(a)**). Those stresses generate clustering of aftershocks near the fault and a systematic variation of the decay exponent p by distance from the earthquake rupture (**Figure 8(b)**). Finally, time-dependent relaxation of an earthquake stress step, perhaps fault creep, can result in aftershock decay with $p > 1$. For example, eqn [34] with values set to relax the stress step with the logarithm of time following an earthquake produces aftershock decay with $p > 1$ (Dieterich, 1994).

4.04.5.3 Foreshocks

Faults with rate- and state-dependent properties appear capable of generating foreshocks by two, somewhat different, mechanisms (Dieterich and Kilgore, 1996a). The first mechanism assumes that mainshocks preceded by foreshocks are in essence aftershocks that happen to be larger than the prior earthquake. With this model, aftershocks are presumed to obey the Gutenberg–Richter frequency distribution of earthquake magnitudes. Hence, there is a finite probability that any aftershock will exceed the magnitude of the prior mainshock. The probability of this is determined by the rates of aftershocks, which are computed with eqn [32] using the stress changes around the region of the prior earthquake, together with the frequency distribution of earthquake magnitudes. The second mechanism assumes foreshocks are driven by mainshock nucleation. In this case, local stressing is proportional to slip speed during mainshock nucleation and accelerates with time (eqns [17] and [20]). The accelerating rate of stressing causes locally accelerating rates of seismicity. For the second mechanism to operate, the dimensions of the mainshock nucleation zone must be significantly greater than the foreshock nucleation sources so that (1) mainshock nucleation is able to drive the foreshock sequences and (2) the region of stress perturbation is of sufficient size to yield an appropriate probability of triggering foreshock. Simple models of each process can be fit to data for foreshock–mainshock pairs (Dieterich and Kilgore, 1996a). However, there is presently no strong evidence to support, or reject, the requirement of the second mechanism that nucleation zones of mainshocks are larger than the nucleation zones of foreshocks.

Numerical simulations of seismicity that implicitly incorporate the ingredients of the first mechanism (rate- and state-dependent nucleation, stress transfer among simulated events, and development of magnitude/frequency distributions for the simulated earthquakes) spontaneously generate foreshocks (Dieterich, 1995; Ziv, 2003). This suggests that at least some foreshocks originate by this mechanism. In a study of foreshocks in the Western United States, Abercrombie and Mori (1995) found decreasing frequency of foreshocks with mainshock depth, which indicates that increasing normal stress inhibits foreshocks. This result is consistent with the normal stress dependence of the solutions in **Table 4** and suggests a triggered origin for foreshock processes.

4.04.5.4 Earthquake Triggering

Triggering of earthquakes by stress transients (principally tides and seismic waves) is a large topic with many unresolved questions. It seems quite reasonable that tidal stresses should have a significant imprint on earthquake occurrence because the stress changes that occur during a tidal cycle of a few hours are generally comparable in magnitude to months of tectonic stressing in seismically active regions. Similarly, large earthquakes produce stress transients that are comparable with decades of tectonic stressing. Although tidal loading (Cochran et al., 2004) and seismic waves (Prejean et al., 2004; West et al., 2005) are known to directly trigger earthquakes, observations show that triggering effects are surprisingly subtle. This can be ascribed to nucleation processes, which are time-dependent, and become

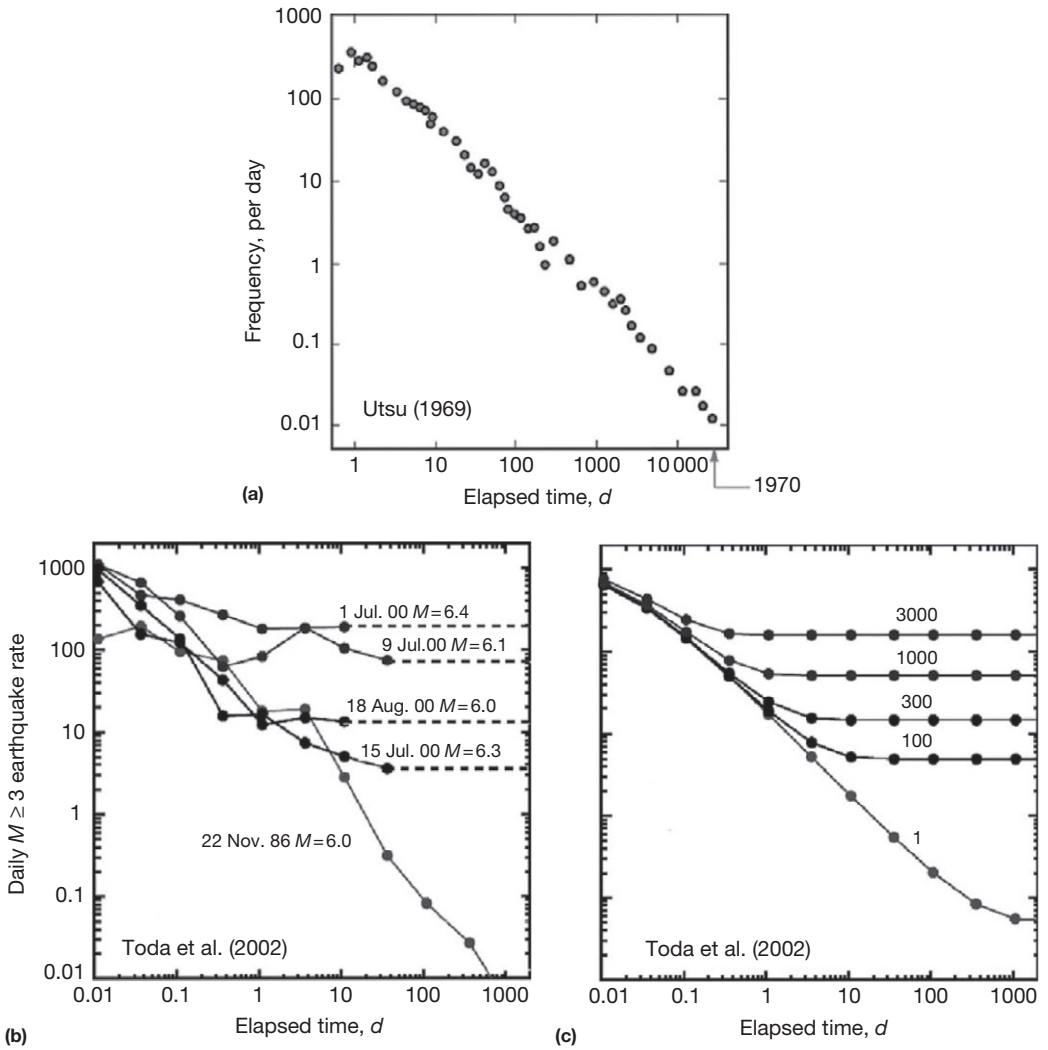


Figure 7 Duration of aftershocks. (a) Aftershocks of the $M \sim 8$, Nobi, Japan, earthquake of 1891 continued at least to 1969 (Utsu, 1969) with no suggestion of returning to a background level after 79 years. Estimated recurrence time for this earthquake is 5000–20 000 years (Tsukuda, 1983), which indicates very low stressing rates. (b) Ultrashort-duration aftershocks of 1–10 days following $M \sim 6$ earthquake that occurred during the intense Izu Islands earthquake swarm of 2000 (Toda et al., 2002). (c) Predicted aftershock decay as a function of the stressing rate for the Izu Islands swarm from eqn [32]. Estimated stressing rates during the large intrusion that caused the swarm were as much as 3000 times larger than the background tectonic stressing rate (indicated on curves).

increasingly insensitive to stress perturbations as the time to instability decreases (Section 4.04.4). The following discussion reviews some characteristics of, and constraints on, earthquake triggering, within the framework of rate- and state-dependent friction.

Constitutive formulations for earthquake rates (eqns [24] and [25]) indicate that earthquake rates will fluctuate in response to tectonic stressing with a superimposed periodic component. Equations [24] and [25] may be evaluated numerically for oscillating stresses using the methods outlined in Section 4.04.5.1. However, useful analytic solutions are readily obtained for conditions of steady-state seismicity $\left(\gamma_{ss} = 1/\dot{S}\right)$, which is perturbed by periodic stresses. For these conditions, time-dependent evolution of γ becomes negligible during stress oscillations with periods less than $\sim 0.1\ t_c$ where

$$t_c = \frac{a\sigma}{\dot{S}} \exp\left(\frac{-\Delta S}{a\sigma}\right) \qquad [37]$$

Tectonic stressing with perturbation by earth tides, ocean tidal loading, and seismic waves falls well below this limit. Hence, evolution with time is negligible and only the direct stress effect as given in eqns [28] and [30] operates on γ. For periodic stressing of the form $S(\phi) = C\cos\phi$, where ϕ is the phase angle and C is the amplitude of the cyclic stress component,

$$\gamma(\phi) = \frac{1}{\dot{S}} \exp\left(\frac{-C\ \cos\phi}{a\sigma}\right) \qquad [38]$$

where $\dot{S}$ is the tectonic stressing rate. This assumes steady-state seismicity $\left(\gamma_{ss} = 1/\dot{S}\right)$ at $\phi = \pi/2,\ \pi/2$. Setting $\dot{S}_r$ and r equal to

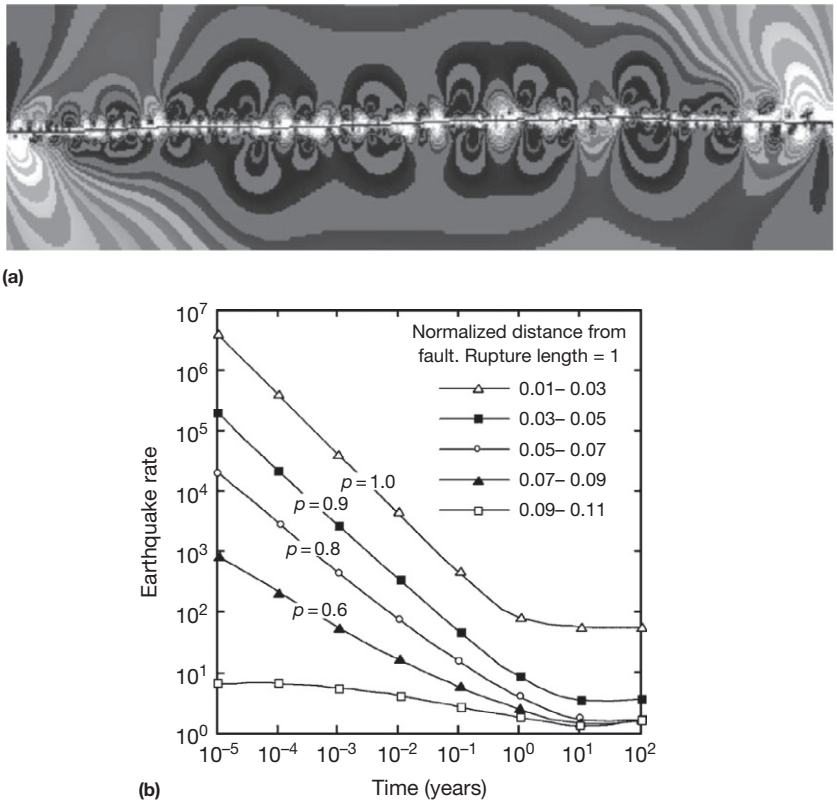


Figure 8 Coulomb stresses and aftershocks near a nonplanar fault (Dieterich, 2005). (a) Boundary element calculation of Coulomb stresses near a fault with random fractal roughness. Light shades indicate areas of stress increase and dark shaded area of stress decrease. (b) Example of calculation of aftershocks as a function of distance from a fractal fault showing spatial variation of aftershock decay exponent p.

the tectonic stressing rate and background steady-state earthquake rate, respectively, gives

$$R(\phi) = r\exp\left(\frac{C\cos\phi}{a\sigma}\right) \qquad [39]$$

Figure 9 shows a fit of this equation to the data of Cochran et al. (2004) for earth and ocean tidal loading.

Distant triggering of earthquakes was noted at the time of the M 7.3 Landers earthquake of 1992 (Hill et al., 1993). Direct triggering of local earthquakes during the passage of the wave train has been widely reported for a number of earthquakes including the 3 November 2002 M 7.9 Denali, Alaska, earthquake where triggering was observed at Mt. Rainier, Washington; Geysers geothermal field, California; Mammoth Mountain/Long Valley Caldera, California; and Coso Geothermal Field, California (Prejean et al., 2004). Some other examples of direct trigging include the 22 December 2003 M 6.5 San Simeon, California, earthquake (Hardebeck et al., 2004); the 6 December 1999 M 7.0 Karluk Lake, Alaska, earthquake (Power et al., 2001); and the 26 December 2004 M 9.3 Sumatra earthquake (West et al., 2005).

The triggering of earthquakes during the passage of stress waves may be evaluated in a similar manner to that used for analysis of tidal stressing because the stressing conditions for seismic waves fall well below the limit where evolution of γ with time is important (eqn [37]); that is, eqns [28] and [30] may be directly applied to follow the change of γ with stress.

Figure 10 illustrates a calculation for triggering by an idealized surface wave. The dependence of triggering on amplitude of stress waves is quite nonlinear, and stress amplitudes of at least $(\Delta S/a\sigma) > 3$ appear necessary to achieve significant levels of triggering. Generic values of $a\sigma$ from aftershock studies (Dieterich, 1994) are in the range 0.1–0.2 MPa, which indicates that somewhat high stress amplitudes, in excess of 0.3 MPa, are required for triggering by short-duration transients. Remote triggering appears to occur preferentially in geothermal and volcanic areas (Hill et al., 1993; Prejean et al., 2004; Stark and Davis, 1996; West et al., 2005) where fluid pressures may be anomalously high and where earthquakes are generally quite shallow – these would reduce the effective normal stress in $(\Delta S/a\sigma)$ permitting a comparable decrease in the required stress-triggering stress ΔS. Additionally, the passage of seismic waves may lead to significant amplification of pore-pressure transients (Brodsky et al., 2003), which would amplify fluctuations in the effective normal component in the stress transient.

Remotely triggered earthquakes often continue for days or weeks after the passage of seismic waves through a region, with the elevated rates decaying in an aftershock-like manner (Brodsky, 2006; Hill et al., 1993). Diverse mechanisms have been proposed to explain continuing elevated rates of seismicity. In the context of rate–state friction, no viable mechanisms have been identified that would explain sustained seismicity as arising from seismic shaking alone. However, aftershock-like

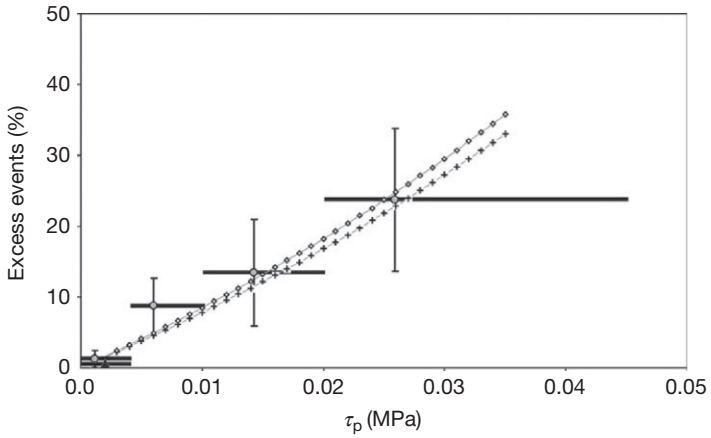


Figure 9 Extra earthquakes N_{ex} occurring at times of positive tidal stressing plotted against the tidal amplitudes τ_p (Cochran et al., 2004). The quantity N_{ex} is the number of earthquakes in excess of the background rate that occur during intervals when stress exceeds the average value, expressed as a percentage of the background earthquakes. The solid triangle near the origin is for the California strike-slip earthquakes. The gray circles are for shallow global thrust earthquakes. The stress amplitudes are mainly driven by ocean tidal loading. The slightly lower curve is a fit to the data using the rate–state eqn [39]. The upper curve is a fit based on a stress corrosion model.

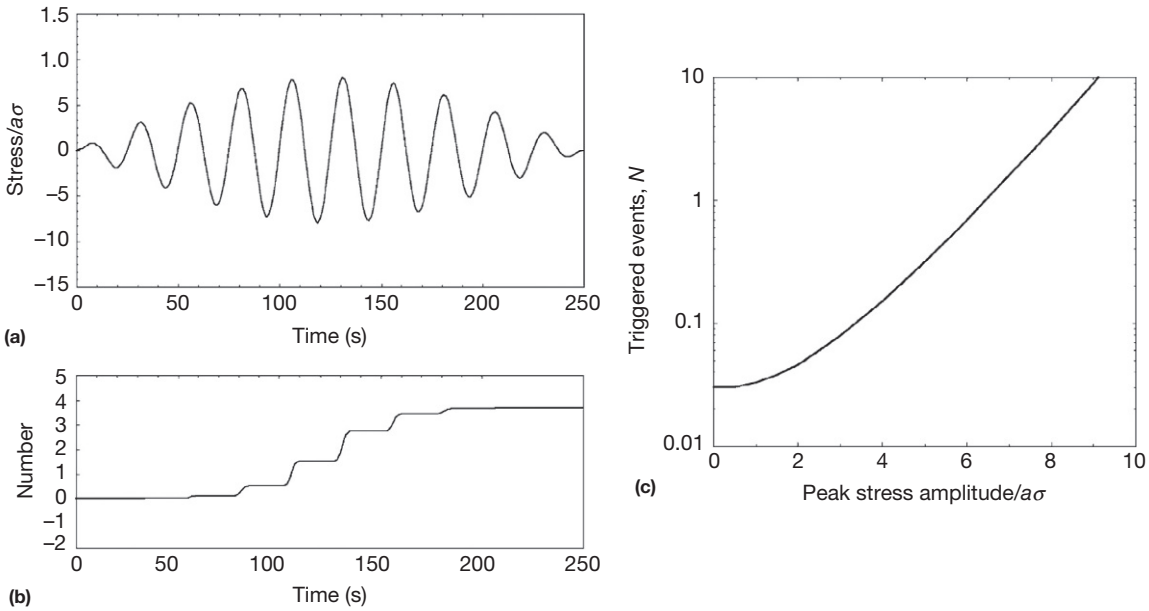


Figure 10 Example of a simulation of dynamic stress triggering. (a) Stresses from an idealized surface wave with peak amplitude of $8a\sigma$. (b) Triggered events calculated with eqns [24] and [30] using (a) as an input. The background rate is set at 10/day. (c) Dependence of number of triggered earthquake on peak amplitude. All parameters are held fixed except for the amplitude of the stress oscillations. Black circle denotes conditions for (a) and (b).

decay of the elevated seismicity suggests that sustained triggering may simply represent an aftershock-like response to a local stress event, triggered by the seismic waves, as described by stress step equations of **Table 4**. Brodsky (2006) presented evidence in support of a proposal by Hough and Kanamori (2002) that continuing earthquakes are normal aftershocks to earthquakes directly triggered by the seismic waves. Other possible sources for triggered local stress events include aseismic fault creep (Anderson et al., 1994; Bodin and Gomber, 1994), fluid pressure steps caused by destruction of flow barriers (Brodsky et al., 2003; Hill et al., 1993), and various adjustments in magmatic systems (Hill et al., 1993).

4.04.6 Stress Changes Estimated from Earthquake Rates

Stress changes in the Earth's crust are generally estimated from model calculations that use near-surface deformation as an observational constraint. The widespread correlation of changes

of earthquake activity with stress (Harrts et al., 1995; King et al., 1994; Reasenberg and Simpson, 1992; Stein, 1999) suggests that stress changes might be calculated from changes of earthquake rates. This possibility has considerable appeal, because seismicity data are routinely collected and have good spatial and temporal resolution. The time and stress dependencies contained in the formulation for earthquake rate, eqns [24] and [25], appear to be consistent with a number of earthquake phenomena over a range of timescales and stressing conditions. The effectiveness of this formulation in forward modeling of earthquake phenomena suggests that it can be used to estimate stress changes from earthquake-rate data.

Recently, a method to do this has been implemented and applied to stress changes at Kīlauea volcano, Hawaii (Dieterich et al., 2000, 2003). It employs observed seismicity rates as a function of time $R(t)$ to directly calculate γ as a function of time. From eqn [24], it follows that

$$\gamma(t) = \frac{r}{R(t)\dot{S}_r} \qquad [40]$$

The differential eqn [25] for γ is then solved for Coulomb stress change as a function of time. The method yields stresses that drive the earthquake process. As such, it is distinct from other seismological methods that yield measures of stress changes resulting from earthquakes.

Figure 11 illustrates some stress-change computations from Kīlauea volcano. Initial results for the stress calculations appear promising – the magnitudes and patterns agree with other estimates of stress change, including those obtained from boundary element modeling of surface deformation observations. The sequence of stress changes shown in Figure 11 indicates a pattern of progressive stress transfer following a dike intrusion event at the onset of the Puu Ōō eruption in 1983. Figure 12 shows a recurring pattern of stress accumulation and release associated with repeating $M \sim 5$ earthquakes on the south flank of Kīlauea volcano.

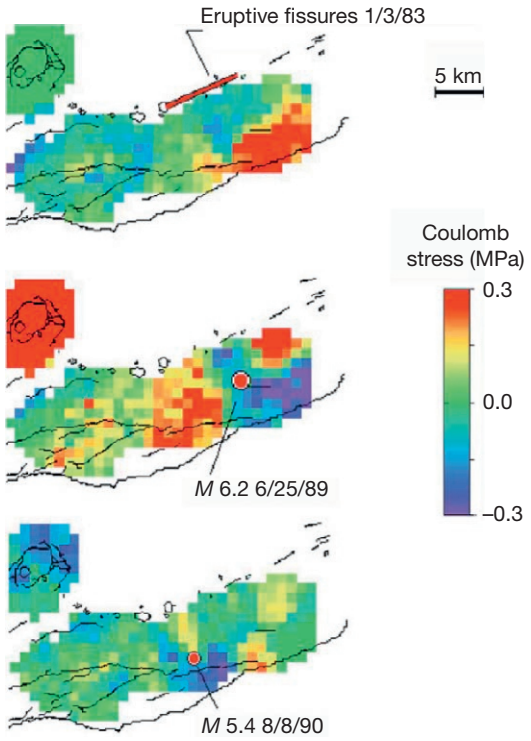


Figure 11 Coulomb stress interactions at Kīlauea volcano, Hawaii, calculated from changes of earthquake rates. Kīlauea caldera is at the upper left of each image. A pattern of progressive stress transfer began with the dike intrusion and eruption at the beginning of 1983 (marked by the eruptive fissures along the southeast rift). The intrusion opened the upper portion of the rift zone by about 2 m, which stressed a portion of the south flank of Kīlauea (red). The region stressed by the intrusion experienced a large earthquake in 1989, which appeared to largely relax the intrusion stress but resulted in stress increase in the region to the southwest. Fourteen months later in 1990, an M 5.4 earthquake occurred in this area. Data from Dieterich JH, Cayol V, and Okubo P (2003) Stress changes before and during the Puu Oo–Kupaianaha eruption. In Heliker C, Donald AS, and Taeko J (eds.) *US Geological Survey Professional Paper 1676: The Puu Oo–Kupaianaha Eruption of Kilauea Volcano, Hawaii: The First 20 Years*, pp. 187–202. Reston, VA: US Geological Survey.

4.04.7 Conclusions and Future Directions

The rate- and state-dependent representation of fault friction has proved to be effective both in characterizing frictional properties in laboratory experiments and in modeling diverse fault-slip and earthquake phenomena. To date, most modeling efforts with rate–state friction have been based on simple spring–slider models and 2-D analyses with planar faults. Faults and fault systems are intrinsically 3-D structures with exquisitely complex geometries. One may speculate that new processes and deeper insights will emerge from fully 3-D analyses – particularly so when those models incorporate the geometric characteristics of faults in nature.

From an experimental standpoint, there is a considerable need for characterizing fault constitutive properties with a variety of lithologies appropriate for subducting sediments, and deep fault zone, and over a wider range of pressure, temperature, and slip rates. For example, there are very few experiments that use the materials and pressure–temperature conditions appropriate for deep fault creep, including the evolving lithologies of subducting plates. The detailed characteristics of frictional properties under these conditions are likely to be of critical importance for understanding and modeling the deep aseismic strain events, which increasingly appear to be characteristic of deep fault slip. Also, slip of nonplanar faults and fault systems necessarily involves coupled interactions among slip, normal stress, and state. Current evolution relations for evolution of state with changes of normal stress are based on a small data set.

Finally, it is noted again that the rate–state evolution equations are entirely empirical. To some extent, the micromechanical processes that give rise to rate–state friction have been characterized. However, we currently lack a comprehensive theoretical framework for deriving the functional dependencies of state evolution from more fundamental parameters – this limits our ability to extrapolate fault properties to conditions that are difficult to access in laboratory experiments.

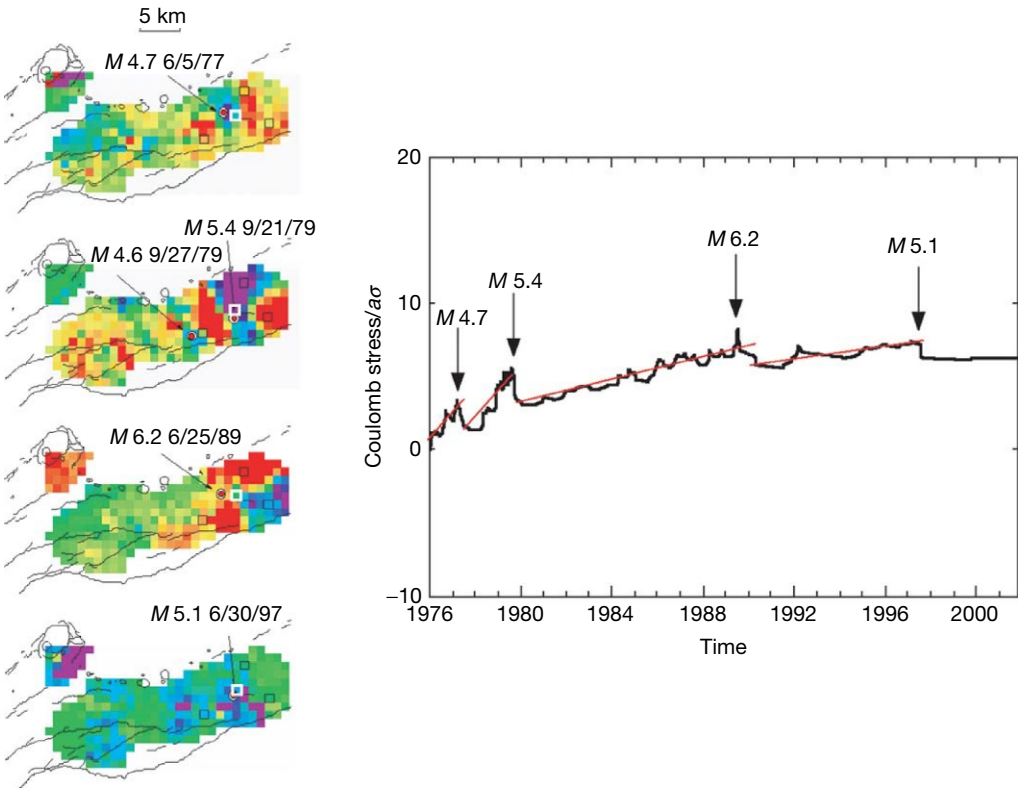


Figure 12 Coulomb stress changes around an area of repeating earthquakes, calculated from changes of earthquake rates. This is the same region shown in **Figure 11**. Left panels show Coulomb stress changes at the time of four earthquakes that affected the same area. In every case, there is a region of decreased stress (stress shadows) near the hypocenters. Right panel shows stress versus time computed from earthquake-rate change at the point marked by the white square in the bottom panel. It shows repeating pattern of progressive stress increase followed by stress drop at the time of the earthquakes. Data from Dieterich JH, Cayol V, and Okubo P (2003) Stress changes before and during the Puu Oo–Kupianaha eruption. In: Heliker C, Donald AS, and Taeko J (eds.) *US Geological Survey Professional Paper 1676: The Puu Oo–Kupaianaha Eruption of Kilauea Volcano, Hawaii: The First 20 Years*, pp. 187–202. Reston, VA: US Geological Survey.

References

Abercrombie RE (1995) Earthquake source scaling relationships from –1 to 5ML using seismograms recorded at 2.5 km depth. *Journal of Geophysical Research* 100: 24015–24036.

Abercrombie RE and Mori J (1995) Occurrence patterns of foreshocks to large earthquakes in the Western United States. *Nature* 381: 303–307.

Anderson JG, Brune JN, Louie J, et al. (1994) Seismicity in the western Great Basin apparently triggered by the Landers, California, earthquake, June 28, 1992. *Bulletin of the Seismological Society of America* 84: 863–891.

Beeler NM, Tullis TE, and Weeks JD (1994) The roles of time and displacement in the evolution effect in rock friction. *Geophysical Research Letters* 21: 1987–1990.

Belardinelli ME, Bizzari A, and Cocco M (2003) Earthquake triggering by static and dynamic stress changes. *Journal of Geophysical Research* 108(B3): 2135. http://dx.doi.org/10.1029/2002JB001779.

Beroza GC and Jordan TH (1990) Searching for slow and silent earthquakes using free oscillations. *Journal of Geophysical Research* 95: 2485–2510.

Blanpied ML, Lockner DA, and Byerlee JD (1991) Fault stability inferred from granite sliding experiments at hydrothermal conditions. *Geophysical Research Letters* 18: 609–612.

Blanpied ML, Lockner DA, and Byerlee JD (1995) Frictional slip of granite at hydrothermal conditions. *Journal of Geophysical Research* 100: 13045–13064.

Blanpied ML and Tullis TE (1986) The stability and behavior of a system with two state variable constitutive law. *Pure and Applied Geophysics* 124: 415–444.

Bodin P and Gomber J (1994) Triggered seismicity and deformation between the Landers, California, and Little Skull Mountain, Nevada earthquakes. *Bulletin of the Seismological Society of America* 84: 835–843.

Brodsky EE (2006) Long-range triggered earthquakes that continue after the wave train passes. *Geophysical Research Letters* 33: L15313. http://dx.doi.org/10.1029/2006GL026605.

Brodsky EE, Roeloffs E, Woodcock D, Gall I, and Manga M (2003) A mechanism for sustained groundwater pressure changes induced by distant earthquakes. *Journal of Geophysical Research* 108: 2390. http://dx.doi.org/10.1029/2002JB002321.

Brooks BA, Foster JH, Bevis M, Fazer LN, Wolfe CJ, and Behn M (2006) Periodic slow earthquakes on the flank of Kilauea Volcano Hawaii. *Earth and Planetary Science Letters* 246: 207–216.

Bucknam RC, Plafker G, and Sharp RV (1978) Fault movement (afterslip) following the Guatemala earthquake of February 4, 1976. *Geology* 6: 170–173.

Burford RO (1976) Continued slip on the Coyote Creek fault after the Borrego Mountain earthquake, in The Borrego Mountain earthquake of April 9, 1968. *US Geological Survey Professional Paper* 787: 105–111.

Bürgmann R, Ergintav S, Segall P, et al. (2002) Time-dependent distributed afterslip on and deep below the Izmit earthquake rupture (in The Izmit, Turkey, earthquake of 17 August 1999, Toksoz). *Bulletin of the Seismological Society of America* 92(1): 126–137.

Byerlee JD (1978) Friction of rocks. *Pure and Applied Geophysics* 116: 615–626.

Cervelli P, Segall P, Johnson K, Lisowski M, and Miklius A (2002) Sudden aseismic fault slip on the south flank of Kilauea Volcano. *Nature* 415: 1014–1018.

Cochran ES, Vidale JE, and Tanaka S (2004) Earth tides can trigger shallow thrust fault earthquakes. *Science* 306: 1164–1166.

Crescentini L, Amoruso A, and Scarpa R (1999) Constraints on slow earthquake dynamics from a swarm in central Italy. *Science* 286: 2132–2134.

Dieterich JH (1972) Time-dependent friction in rocks. *Journal of Geophysical Research* 77: 3690–3697.

Dieterich JH (1979) Modelling of rock friction. Part 1: Experimental results and constitutive equations. *Journal of Geophysical Research* 84(B5): 2161–2168.

Dieterich JH (1981) Constitutive properties of faults with simulated gouge. In: Carter NL, Friedman M, Logan JM, and Stearns DW (eds.) *Monograph 24: Mechanical Behavior of Crustal Rocks*, pp. 103–120. Washington, DC: American Geophysical Union.

Dieterich JH (1986) A model for the nucleation of earthquake slip. In: Das S, Boatwright J, and Scholz C (eds.) *Geophysical Monograph 37: Earthquake Source Mechanics*, pp. 37–47. Washington, DC: American Geophysical Union.
Dieterich JH (1987) Nucleation and triggering of earthquake slip: Effect of periodic stresses. *Tectonophysics* 144: 127–139.
Dieterich JH (1992) Earthquake nucleation on faults with rate-and state-dependent friction. *Tectonophysics* 211: 115–134.
Dieterich JH (1994) A constitutive law for rate of earthquake production and its application to earthquake clustering. *Journal of Geophysical Research* 99: 2601–2618.
Dieterich JH (1995) Earthquake simulations with time-dependent nucleation and long-range interactions. *Journal of Nonlinear Processes in Geophysics* 2: 109–120.
Dieterich JH (2005) Role of stress relaxation in slip of geometrically complex faults. *Abstract, EOS, Transactions American Geophysical Union,* (Fall Meeting Supplement) NG31A-04.
Dieterich JH, Cayol V, and Okubo P (2000) The use of earthquake rate changes as a stress meter at Kilauea Volcano. *Nature* 408: 457–460.
Dieterich JH, Cayol V, and Okubo P (2003) Stress changes before and during the Puu Oo–Kupianaha eruption. In: Heliker C, Donald AS, and Taeko J (eds.) *US Geological Survey Professional Paper 1676: The Puu Oo–Kupaianaha Eruption of Kilauea Volcano, Hawaii: The First 20 Years*, pp. 187–202. Reston, VA: US Geological Survey.
Dieterich JH and Conrad G (1984) Effect of humidity on time-and velocity-dependent friction in rocks. *Journal of Geophysical Research* 89: 4196–4202.
Dieterich JH and Kilgore BD (1994) Direct observation of frictional contacts: New insights for sliding memory effects. *Pure and Applied Geophysics* 143: 283–302.
Dieterich JH and Kilgore B (1996a) Implications of fault constitutive properties for earthquake prediction. *Proceedings of the National Academy of Sciences United States of America* 93: 3787–3794.
Dieterich JH and Kilgore B (1996b) Imaging surface contacts; power law contact distributions and contact stresses in quartz, calcite, glass, and acrylic plastic. *Tectonophysics* 256: 219–239.
Dieterich JH and Linker MF (1992) Stability of fault slip under condition of variable normal stress. *Geophysical Research Letters* 19: 1691–1694.
Dragert H, Wang K, and James TS (2001) A silent slip event on the deeper Cascadia subduction interface. *Science* 292: 1525–1528.
Dvorak JJ, Okamura AT, English TT, et al. (1986) Mechanical response of the south flank of Kilauea Volcano, Hawaii, to intrusive events along the rift systems. *Tectonophysics* 124: 193–209.
Evans B and Goetz C (1979) The temperature variation of hardness of olivine and its implication for polycrystalline yield stress. *Journal of Geophysical Research* 84: 5505–5524.
Freymueller JT, Hreinsdóttir S, Zweck C, and Haeussler PJ (2002) The 1998–2002 deep megathrust slip event, Alaska. Eos Trans. AGU Fall. *Meet. Suppl*. abstract G61A–0972.
Fujii Y (1993) Vertical crustal movement in the Boso peninsula, South Kanto, Japan, as deduced from the adjustment of a geodetic network with signals. *Tectonophysics* 218: 309–322.
Gladwin MT, Gwyther RL, Hart RHG, and Breckenridge KS (1994) Measurements of the strain field associated with episodic creep events on the San Andreas fault at San Juan Bautista, California. *Journal of Geophysical Research* 99(B3): 4559–4566.
Goldsby DL, Rar A, Pharr GM, and Tullis TE (2004) Nanoindentation creep of quartz with implications for rate- and state-variable friction laws relevant to earthquake mechanics. *Journal of Materials Research* 19(1): 357–365.
Gomberg J, Beeler N, and Blanpied M (2000) On rate-state and Coulomb failure models. *Journal of Geophysical Research* 105: 7557–7871.
Gomberg J, Beeler NM, Blanpied ML, and Bodin P (1998) Earthquake triggering by transient and static deformations. *Journal of Geophysical Research* 103: 24411–24426.
Gomberg J, Blanpied ML, and Beeler NM (1997) Transient triggering of near and distant earthquakes. *Bulletin of the Seismological Society of America* 87: 294–309.
Gu J, Rice JR, Ruina AL, and Tse ST (1984) Slip motion and stability of a single degree of freedom elastic system rate and state dependent friction. *Journal of the Mechanics and Physics of Solids* 32: 167–196.
Hardebeck JL, Boatwright J, Dreger D, et al. (2004) Preliminary report on the 22 December 2003 M6.5 San Simeon, California, earthquake. *Seismological Research Letters* 75: 155–172.
Harrts RA, Simpson R, and Reasenberg PA (1995) Influence of static stress changes on earthquake locations in southern California. *Nature* 375: 221–224.
Hill DP, Reasenberg PA, Michael A, et al. (1993) Seismicity remotely triggered by the Magnitude 7.3 Landers, California, earthquake. *Science* 260: 1617–1623.
Hirose H, Hirahara K, Kimata F, Fujii N, and Miyazaki S (1999) A slow thrust slip event following the two 1996 Hyuganada earthquakes beneath the Bungo Channel, southwest Japan. *Geophysical Research Letters* 26: 3237–3240.
Hough S and Kanamori H (2002) Source properties of earthquakes near the Salton Sea triggered by the 16 October 1999 M 7.1 Hector Mine, California, earthquake. *Bulletin of the Seismological Society of America* 92: 1281–1289.
Hsu Y-J, Bechor N, Segall P, Yu S-B, Kuo L-C, and Ma K-F (2002) Rapid afterslip following the 1999 Chi-Chi, Taiwan Earthquake. *Geophysical Research Letters* 29. http://dx.doi.org/10.1029/2002GL014967.
Hutton W, DeMets C, Sánchez O, Suárez G, and Stock J (2002) Slip kinematics and dynamics during and after the 1995 October 9 $Mw = 8.0$ Colima–Jalisco earthquake, Mexico, from GPS geodetic constraints. *Geophysical Journal International* 146: 637.
Kanamori H and Hauksson E (1992) A slow earthquake in the Santa Maria Basin, California. *Bulletin of the Seismological Society of America* 82: 2087–2096.
Kato N and Hirasawa T (1997) A numerical study on seismic coupling along subduction zones using a laboratory-derived friction law. *Physics of the Earth and Planetary Interiors* 102: 51–68.
Kawasaki I (2004) Silent earthquakes occurring in a stable–unstable transition zone and implications for earthquake prediction. *Earth, Planets and Space* 56: 813–821.
Kilgore BD, Blanpied ML, and Dieterich JH (1993) Velocity-dependent friction of granite over a wide range of conditions. *Geophysical Research Letters* 20: 903–906.
King GCP, Stein RS, and Lin J (1994) Static stress changes and the triggering of earthquakes. *Bulletin of the Seismological Society of America* 84: 935–953.
Kodaira S, Lidaka T, Kato A, Park J-O, Iwasaki T, and Kaneda Y (2004) High pore fluid pressure may cause silent slip in the Nankai Trough. *Science* 304: 1295–1298.
Kostoglodov V, Singh SK, Santiago JA, and Franco SI (2003) A large silent earthquake in the Guerrero seismic gap, Mexico. *Geophysical Research Letters* 30: 1807. http://dx.doi.org/10.1029/2003GL017219.
Langbein JO (1981) An interpretation of episodic slip on the Calaveras fault near Hollister, California. *Journal of Geophysical Research* 86: 4941–4948.
Lapusta N and Rice JR (2003) Nucleation and early seismic propagation of small and large events in a crustal earthquake model. *Journal of Geophysical Research* 108(B4): 2205. http://dx.doi.org/10.1029/2001JB000793.
Lapusta N, Rice JR, Ben-Zion Y, and Zheng G (2000) Elastodynamic analysis for slow tectonic loading with spontaneous rupture episodes on faults with rate- and state-dependent friction. *Journal of Geophysical Research* 105(B10): 23765–23789.
Linde AT and Sacks IS (2002) Slow earthquakes and great earthquakes along the Nankai trough. *Earth and Planetary Science Letters* 203: 265–275.
Linker MF and Dieterich JH (1992) Effects of variable normal stress on rock friction: Observations and constitutive relations. *Journal of Geophysical Research* 97: 4923–4940.
Liu Y and Rice JR (2005) Aseismic slip transients emerge spontaneously in three-dimensional rate and state modeling of subduction earthquake sequences. *Journal of Geophysical Research* 110: B08307. http://dx.doi.org/10.1029/2004JB003424.
Lowry AR, Larson KM, Kostoglodov V, and Bilham R (2001) Transient fault slip in Guerrero, southern Mexico. *Geophysical Research Letters* 28: 3753–3756.
Marone C (1998) Laboratory-derived friction laws and their application to seismic faulting. *Annual Review of Earth and Planetary Sciences* 26: 643–696.
Marone C and Kilgore B (1993) Scaling of the critical slip distance for seismic faulting with shear strain in fault zones. *Nature* 362: 618–621.
Marone C, Raleigh CB, and Scholz CH (1990) Frictional behavior and constitutive modeling of simulated fault gouge. *Journal of Geophysical Research* 95: 7007–7025.
Marone CJ, Scholz CH, and Bilham R (1991) On the mechanics of earthquake afterslip. *Journal of Geophysical Research* 96(B5): 8441–8452.
McHugh S and Johnston MJS (1976) Short period non-seismic tilt perturbations and their relation to episodic slip on the San Andreas fault. *Journal of Geophysical Research* 81: 6341–6346.
Miller MM, Melbourne TI, Johnson DJ, and Sumner WQ (2002) Periodic slow earthquakes from the Cascadia Subduction Zone. *Science* 295: 2423.
Miyazaki S, Segall SP, Fukuda J, and Kato T (2004) Space time distribution of afterslip following the 2003 Tokachi-oki earthquake: Implications for variations in fault zone frictional properties. *Geophysical Research Letters* 31. http://dx.doi.org/10.1029/2003GL019410.
Nason RD and Weertman J (1973) A dislocation theory analysis of fault creep events. *Journal of Geophysical Research* 78: 7745–7751.
Okubo PG (1989) Dynamic rupture modeling with laboratory-derived constitutive relations. *Journal of Geophysical Research* 94: 12321–12335.
Okubo PG and Dieterich JH (1986) State variable fault constitutive relations for dynamic slip. In: Das S, Boatwright J, and Scholz C (eds.) *Monograph Series 37: Earthquake Source Mechanics*, pp. 25–35. Washington, DC: AGU.

Ozawa S, Murakami M, Kaidzu M, et al. (2002) Detection and monitoring of ongoing aseismic slip in the Tokai region, central Japan. *Science* 298: 1009–1012.
Perrin G, Rice JR, and Zheng G (1995) Self-healing slip pulse on a frictional surface. *Journal of the Mechanics and Physics of Solids* 43: 1461–1495.
Pollitz F, Bürgmann R, and Segall P (1998) Joint estimation of afterslip rate and postseismic relaxation following the 1989 Loma Prieta earthquake. *Journal of Geophysical Research* 103: 26975–26992.
Power JA, Moran SC, McNutt SR, Stihler SD, and Sanchez JJ (2001) Seismic response of the Katmai volcanoes to the 6 December 1999 Magnitude 7.0 Karluk Lake earthquake, Alaska. *Bulletin of the Seismological Society of America* 91: 57–63.
Prakash V (1998) Frictional response of sliding interfaces subjected to time varying normal pressures. *Journal of Tribology* 120: 97–102.
Prejean SG, Hill DP, Brodsky EE, et al. (2004) Remotely triggered seismicity on the United States West Coast following the *Mw* 7.9 Denali fault earthquake. *Bulletin of the Seismological Society of America* 94: S348–S359.
Reasenberg PA and Simpson RW (1992) Response of regional seismicity to the static stress change produced by the Loma Prieta earthquake. *Science* 255: 1687–1690.
Rice JR (1983) Constitutive relations for fault slip and earthquake instabilities. *Pure and Applied Geophysics* 121: 443–475.
Rice JR (2006) Heating and weakening of faults during earthquake slip. *Journal of Geophysical Research* 111: BO5311. http://dx.doi.org/10.1029/2005JB004006.
Rice JR and Ruina AL (1983) Stability of steady frictional slipping. *Journal of Applied Mechanics* 50: 343–349.
Rogers G and Dragert H (2003) Episodic tremor and slip on the cascadia subduction zone: The chatter of silent slip. *Science* 300: 1942–1992.
Rubin AM and Ampuero J-P (2005) Earthquake nucleation on (aging) rate and state faults. *Journal of Geophysical Research* 110: B11312. http://dx.doi.org/10.1029/2005JB003686.
Ruina AL (1983) Slip instability and state variable friction laws. *Journal of Geophysical Research* 88: 10359–10370.
Sacks IS, Linde AT, Snoke JA, and Suyehiro S (1981) A slow earthquake sequence following the Izu–Oshima earthquake of 1978. In: Simpson DW and Richards PG (eds.) *Maurice Ewing Series*. *Earthquake Prediction: An International Review*, vol. 4, pp. 617–628. Washington, DC: AGU.
Sagiya T (2004) Interplate coupling in the Kanto district, central Japan, and the Boso Peninsula silent earthquake in May 1996. *Pure and Applied Geophysics* 161: 2327–2342.
Scholz CH and Engelder JT (1976) The role of asperity indentation and ploughing in rock friction, Part I: Asperity creep and stick-slip. *International Journal of Rock Mechanics and Mining Science Geomechanics Abstract* 77: 149–154.
Segall P and Rice JR (1995) Dilatancy, compaction, and slip instability of a fluid-infiltrated fault. *Journal of Geophysical Research* 100(B11): 22155–221121.
Shibazaki B and Iio Y (2003) On the physical mechanism of silent slip events along the deeper part of the seismogenic zone. *Geophysical Research Letters* 30: 1489–1492.
Sleep NH (1995a) Ductile creep compaction and rate and state dependent friction with major fault zones. *Journal of Geophysical Research* 100: 13065–13080.
Sleep NH (1995b) Frictional heating and the stability of rate and state dependent frictional sliding. *Geophysical Research Letters* 22: 2785–2788.
Sleep NH (1997) Application of a unified rate and state dependent theory to the mechanics of fault zones with strain localization. *Journal of Geophysical Research* 102: 2875–2895.
Sleep NH and Blanpied ML (1992) Creep, compaction and the weak rheology of major faults. *Nature* 359: 687–692.
Smith SW and Wyss M (1968) Displacement on the San Andreas fault subsequent to the 1966, Parkfield earthquake. *Bulletin of the Seismological Society of America* 58: 1955–1973.
Stark MA and Davis SD (1996) Remotely triggered microearthquakes at The Geysers geothermal field, California. *Geophysical Research Letters* 23: 945–948.
Stein RS (1999) The role of stress transfer in earthquake occurrence. *Nature* 402: 605–609.
Stesky RM (1978) Rock friction – Effect of confining pressure, temperature, and pore pressure. *Pure and Applied Geophysics* 116: 690–704.
Stuart W (1988) Forecast model for great earthquakes at the Nankai Trough subduction zone. *Pure and Applied Geophysics* 126: 619–641.
Toda S, Stein RS, Reasenberg PA, Dieterich JH, and Yoshida A (1998) Stress transferred by the 1995 Mw = 6.9 Kobe, Japan, shock. *Journal of Geophysical Research* 103: 24543–24565.
Toda S, Stein RS, and Sagiya S (2002) Evidence from the AD 2000 Izu islands earthquake swarm that stressing rate governs seismicity. *Nature* 419: 58–61.
Tse ST and Rice JR (1986) Crustal earthquake instability in relation to the depth variation of frictional slip properties. *Journal of Geophysical Research* 91(B9): 9452–9472.
Tullis TE (1988) Rock friction constitutive behavior from laboratory experiments and its implications for and earthquake prediction field monitoring program. *Pure and Applied Geophysics* 126: 555–588.
Tullis TE and Weeks JD (1986) Constitutive behavior and stability of frictional sliding in granite. *Pure and Applied Geophysics* 124: 383–414.
Utsu T (1969) Aftershocks and earthquake statistics. *Journal of the Faculty of Science, Hokkaido University* 7(3): 129–195.
Walker WW and Demer LJ (1964) Effects of loading duration on indentation hardness. *Transactions of the American Institute of Mining, Metallurgical and Petroleum Engineers* 233: 425–428.
Weeks JD and Tullis TE (1985) Frictional sliding of dolomite: A variation in constitutive behavior. *Journal of Geophysical Research* 90: 7821–7826.
Wennerberg L and Sharp RV (1997) Bulk-friction modeling of afterslip and the modified Omori law. *Tectonophysics* 277: 109–136.
Wesson RL (1988) Dynamics of fault creep. *Journal of Geophysical Research* 93(B8): 8929–8951.
West M, Sanchez JJ, and McNutt SR (2005) Periodically triggered seismicity at Mount Wrangell, Alaska, after the Sumatra earthquake. *Science* 308: 1144–1146.
Westbrook JH and Jorgensen PJ (1965) Indentation creep of solids. *Transactions of the American Institute of Mining, Metallurgical and Petroleum Engineers* 230: 613–614.
Ziv A (2003) Foreshocks, aftershocks and remote triggering in Quasi-static fault models. *Journal of Geophysical Research* 108. http://dx.doi.org/10.1029/2002JB002318.

4.05 The Mechanics of Frictional Healing and Slip Instability During the Seismic Cycle

C Marone and DM Saffer, The Pennsylvania State University, University Park, PA, USA

© 2015 Elsevier B.V. All rights reserved.

4.05.1	**Introduction**	111
4.05.1.1	Application to Faulting	113
4.05.1.2	Shear Failure, Frictional Healing, and Friction Constitutive Laws	114
4.05.1.3	Repetitive Frictional Failure and the Upper Stability Transition from Stable to Unstable Sliding	115
4.05.2	**Laboratory Experiments**	115
4.05.2.1	Transducers, Control Electronics, and Data	115
4.05.2.2	Materials	116
4.05.2.3	Velocity and Frictional Force Measurements	116
4.05.3	**Laboratory Data**	116
4.05.3.1	Experimental Boundary Conditions and Range of Major Parameters	116
4.05.3.2	SHS Tests, Creep, and Static Friction	118
4.05.3.3	Frictional Healing, Creep Consolidation, and Dilation	118
4.05.3.4	Effect of Slip Velocity on Frictional Healing	121
4.05.3.5	Creep, Precursory Slip, and Relocalization of Strain Following Healing	121
4.05.4	**Analysis and Discussion**	123
4.05.4.1	Forward Modeling of Frictional Healing	126
4.05.4.2	Origin of the Loading Rate Effect on Healing	127
4.05.4.3	Recovery of Rate State Friction Parameters from SHS Tests	132
4.05.4.4	Microphysical Explanation of Frictional Healing and Rate Dependence	135
4.05.5	**Conclusions**	135
Acknowledgments		135
References		135

4.05.1 Introduction

Stick-slip frictional motion is a common characteristic of a wide variety of mechanical systems, ranging from wood flooring to automobile brakes, polymer films, and plate tectonic faults (Persson, 1998; Scholz, 2002). Most studies of stick-slip friction focus on the dynamic motion during slip because of its importance in generating vibrations, heat, wear, and, in some cases, seismic radiation. However, the processes that govern friction recovery (healing) and the nucleation of instabilities are also fundamental to understanding earthquakes and the seismic cycle. For example, many types of observations exhibit strong signatures of frictional healing, including (1) the depth frequency distribution of seismicity at subduction megathrusts (**Figure 1**) and (2) the structural evidence of repeated failure – in the form of the deformed margin wedge and faults therein. Geologic and geophysical data record both repeated seismic and aseismic slip on faults and the way faults regain frictional strength after failure.

Subduction zones host the world's largest earthquakes, and it is well known that large earthquakes occur on preexisting, large faults that have experienced multiple cycles of seismic failure (e.g., Scholz, 2002). Moreover, in evaluating the distribution of seismic and aseismic deformation along Earth's major plate bounding faults, including subduction zones, a key unresolved question centers on processes and rock properties that control the upper transition, at a depth of ~5–10 km, from aseismic to seismic slip (**Figure 1**) (e.g., Beeler, 2007; Hyndman et al., 1997; Ikari et al., 2007; Marone and Scholz, 1988; Moore and Saffer, 2001; Saffer and Marone, 2003). To host repeated earthquakes, these structures must exhibit *both* frictional healing *and* frictional and elastic properties that allow instabilities to nucleate. In this sense, the processes of stick-slip frictional failure and frictional strength recovery are intimately related.

Both the rate dependence of friction that affects sliding stability and nucleation and the recovery of frictional strength are governed by processes that can be described by constitutive equations applicable to the complete cycle of stick-slip failure (e.g., Dieterich, 1992, 1994; Gu et al., 1984; Rice and Ruina, 1983; Tse and Rice, 1986). Much progress has been made in developing friction constitutive laws to describe variations in the frictional strength of faults with slip rate, slip history, contact time, normal stress, and temperature (Angevine et al., 1982; Beeler, 2007; Beeler et al., 1994, 2008; Behnsen and Faulkner, 2012; Blanpied and Tullis, 1986; Boettcher and Marone, 2004; Bos and Spiers, 2000, 2002; Carpenter et al., 2011; Chester, 1994; Chester and Higgs, 1992; den Hartog and Spiers, 2013; den Hartog et al., 2012a,b; Dieterich, 1978, 1979, 1981; Dieterich and Kilgore, 1994, 1996a,b; Fredrich and Evans, 1992; Frye and Marone, 2002; Goldsby and Tullis, 2011; Hong and Marone, 2005; Ikari et al., 2011; Kanagawa et al., 2000; Karner and Marone, 1998; Karner and Marone, 2001; King and Marone, 2012; Linker and Dieterich, 1992;

 http://dx.doi.org/10.1016/B978-0-444-53802-4.00092-0

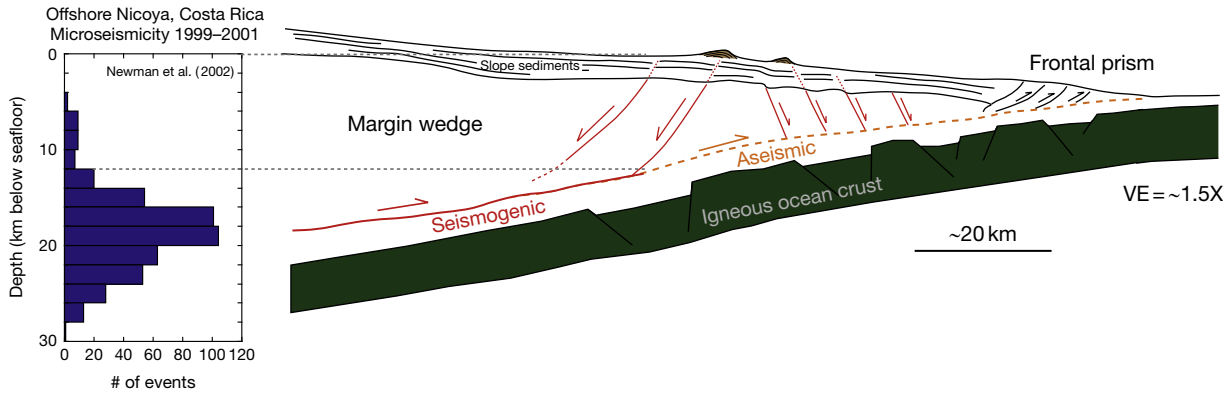


Figure 1 Cross section showing basic structural features of the subduction zone megathrust at Costa Rica. The downgoing slab includes igneous ocean crust overlain by marine sediments; major faults include the décollement (subduction plate boundary) and a series of faults in the upper plate margin wedge. Left panel shows histogram of seismicity from Costa Rica (Newman et al., 2002). Dashed line shows the upper stability transition from aseismic to seismogenic fault behavior, defined as the depth below which 90% of the seismicity occurs.

Marone, 1998a,b; Marone and Saffer, 2007; Marone et al., 2009; Muhuri et al., 2003; Niemeijer and Spiers, 2005, 2006, 2007; Niemeijer et al., 2008, 2010; Ohnaka et al., 1997; Olsen et al., 1998; Paterson, 1995; Perrin et al., 1995; Persson, 1998; Persson, 1998; Reinen et al., 1994; Renard et al., 2012; Rice, 1983; Rice, 1993; Richardson and Marone, 1999; Rubinstein et al., 2004; Rubin, 2008; Ruina, 1983; Saffer and Marone, 2003; Saffer et al., 2001; Scholz, 1998; Segall and Rice, 1995; Shimamoto and Logan, 1981a,b; Tenthorey and Cox, 2006; Tenthorey et al., 2003; Tesei et al., 2012; Tullis, 1988; Yasuhara et al., 2003, 2005; Zheng and Elsworth, 2012, 2013). In particular, the rate dependence of friction and its link to potential unstable sliding have been explored extensively through laboratory experiments, for both natural and synthetic fault materials (e.g., Brown et al., 2003; Carpenter et al., 2011, 2012; Ikari et al., 2007, 2009, 2011; Saffer and Marone, 2003). In contrast, frictional healing has received relatively less attention.

Notwithstanding the considerable progress made in laboratory experiments to illuminate the mechanics of frictional aging (e.g., Dieterich and Kilgore, 1994, 1996b), and in particular recent work at the atomic scale (e.g., Li et al., 2011), we have only a rudimentary understanding of the underlying processes and rates of frictional healing in fault zones. For example, in the context of the simplest friction law, involving only static and kinetic friction, the question arises: how is static friction reset between slip events? For some materials, an understanding is beginning to emerge based on plastic deformation and adhesion at highly stressed contact points (Ben-David and Fineberg, 2011; Ben-David et al., 2010a,b; Bhushan et al., 1995; Dieterich and Kilgore, 1994, 1996b; He et al., 1999; Johnson, 1996, 1997; Li et al., 2011; McLaskey et al., 2012). However, for fault rocks and granular materials, the answer is less clear. Is fault healing a temporally continuous process, or does it operate only below a threshold slip velocity? Do depth variations in frictional healing processes control the updip limit of the seismogenic zone (e.g., **Figure 1**)? How do the properties and compositions of mineral surfaces, pore fluids, and fault gouge affect frictional healing? What role, if any, does healing play in determining the nature of earthquake rupture propagation (e.g., Liu, 2013; Marone, 1998b; Marone et al., 1995; Obara, 2002; Peng and Gomberg;, 2010; Peng et al., 2005; Rubin and Ampuero, 2005; Shelly et al., 2007; Tadokoro and Ando, 2002; Vidale et al., 1994; Zheng and Rice, 1998)? These are complex issues involving a large number of coupled processes, but a better grasp of the mechanics of frictional healing is an important step toward understanding the seismic cycle and the spatiotemporal character of the seismogenic zone.

We focus here on the nature of repetitive stick-slip instability, frictional healing, and the nature of the stability transition from seismic to aseismic deformation in tectonic fault zones, through a review of laboratory experimental results coupled with numerical simulations. We draw in particular from direct shear experiments on simulated fault gouge at normal stresses appropriate for the upper part of the seismogenic zone and for shearing velocities appropriate for earthquake nucleation, in the range of tens to hundreds of micron/s (**Figure 2**). This chapter is in part a review and in part a presentation of new results. The review is extensive but not fully comprehensive, due to time and space constraints, and because we draw more heavily on our own work than is probably fully justified.

The primary purposes of this chapter are (1) to explore the nature of frictional restrengthening and fault healing, with particular emphasis on laboratory experiments focused on the effect of loading rate on frictional healing, (2) to evaluate the role of consolidation and lithification for effecting changes in elastic properties of subduction zone and, in general, fault zone materials, and (3) to investigate the possible relationships between elastic stiffness, frictional healing, and frictional stability. The latter issue in particular has received relatively little attention, compared to other possibilities, as an explanation for the updip limit of the seismogenic zone. We are able to more rigorously consider this explanation using new friction data and measurements of elastic parameters from recent field studies. In particular, the expected changes in frictional behavior and elastic stiffness inferred from laboratory experiments on relevant materials are consistent with the observed transition from stable, aseismic deformation to unstable, seismic slip at depths of ~5–10 km (**Figures 1** and **3**).

In the context of Coulomb failure including cohesion and friction (or internal friction), time-dependent strengthening can occur via increased friction or cohesion. Existing works

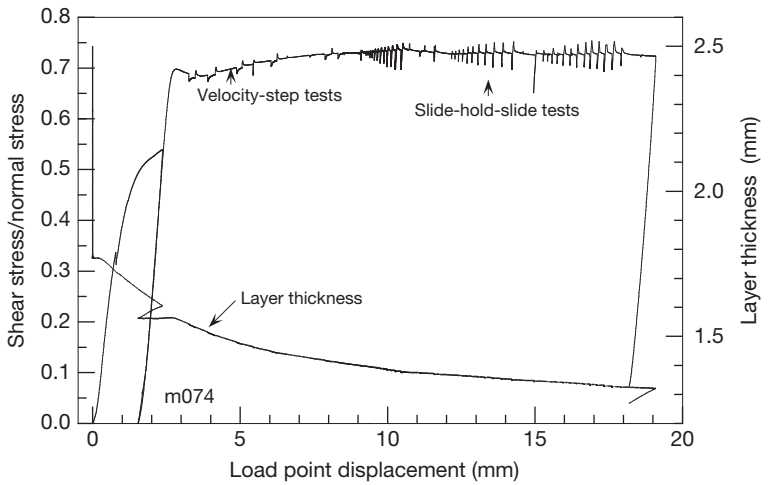


Figure 2 Friction and layer thickness data from a representative experiment. Strain hardening occurs upon initial shearing and is terminated by removing and reapplying shear load. Velocity step tests are used to assess the progression toward a steady-state friction behavior. Three sets of slide-hold-slide (SHS) tests are shown. Layer thickness data show rapid compaction during initial strain hardening and a progressive decrease in compaction rate with continued shear. The roughly linear rate of thinning reached after ≈10 mm of slip corresponds to geometric thinning of the layer. Vertical line at zero displacement shows compaction associated with the application of normal load (from 0 to 25 MPa).

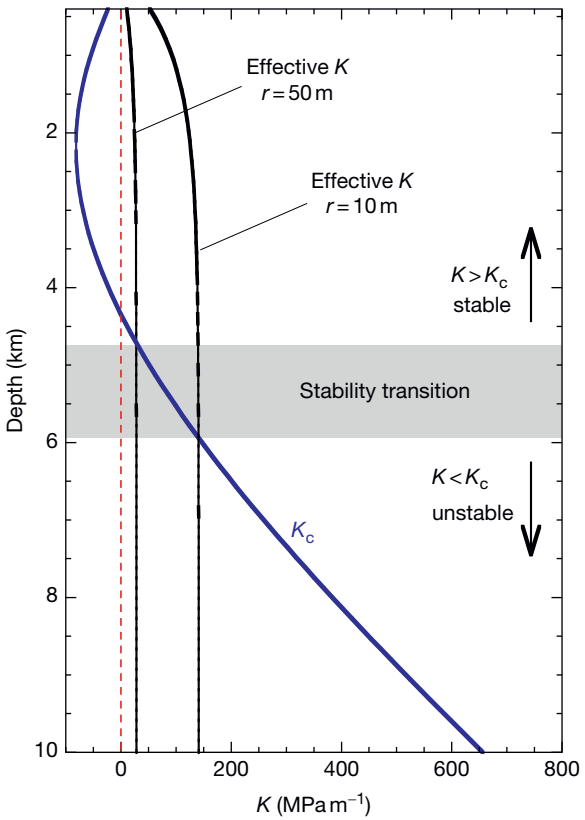


Figure 3 Plot showing K and K_c as a function of depth along the plate interface for a subduction zone example. Rate dependence of friction ($b - a$) is defined by fitting the data of Ikari et al. (2007); D_c is fixed at 30 μm. Effective normal stress is computed assuming pore pressure is 85% of the lithostatic stress (e.g., Davis et al., 1983). The effective stiffness is computed for a circular slip patch of radius 10 and 50 m, respectively, embedded in an elastic half-space. For this range of parameters, a transition from stable to unstable slip would occur where $K < K_c$, at a depth of ~4–6 km.

address these issues to some extent (Karner et al., 1997; Marone et al., 1992; Muhuri et al., 2003), and additional work is needed. We focus here on the normal stress-dependent part of the shear strength, which is the frictional part.

4.05.1.1 Application to Faulting

Seismogenic fault zones are expected to be spatiotemporally complex regions, hosting a range of mineral types, fracture geometries, and fluids as well as a spectrum of slip velocities and strain rates ranging from dynamic to interseismic (Beroza and Ide, 2011; Ide et al., 2007; **Figure 1**). Field observations show that faults contain granular wear materials, referred to as fault gouge, and that these zones can be up to hundreds of meters wide (e.g., Chester and Chester, 1998; Li and Vidale, 2001; Li et al., 1998; Montgomery and Jones, 1992; Unsworth et al., 1997). Ideally, this range of complexity would be included in laboratory experiments designed to study fault zone rheology. However, our current understanding of friction falls well short of that necessary to incorporate the full range of expected complexity within natural fault zones. One problem is that of the sensitivity of slip stability to small variations in frictional properties (Rice and Ruina, 1983). The stability of frictional sliding is often determined by subtle variations of friction representing only a small fraction of the total frictional force, and this appears to apply to a wide range of materials (e.g., Baumberger et al., 1994; Beeler, 2007; Collettini et al., 2011; Dieterich and Kilgore, 1994; McLaskey et al., 2012; Persson, 1998; Scholz, 2002).

Previous work shows that all of the following variables have significant effects on the stability of fault slip: particle size distribution of gouge, particle comminution during shear, shear localization and fabric development, clay content and mineralogy, roughness of the bounding surfaces, net slip, porosity, fluid chemistry, temperature, effective normal stress, and absolute slip velocity (e.g., Beeler et al., 1996; Boulton et al., 2012; den Hartog et al., 2012a,b; Di Toro et al., 2006,

2011; Dieterich and Kilgore, 1994; Marone, 1998a; Niemeijer and Spiers, 2005, 2006; Niemeijer et al., 2011; Noda and Shimamoto, 2009; Saffer et al., 2001; Tesei et al., 2012). Tackling all of these in a single study has proven challenging. Instead, laboratory investigations have generally focused on isolating the effects of a few variables.

4.05.1.2 Shear Failure, Frictional Healing, and Friction Constitutive Laws

Measurements of frictional aging have been available for over 200 years, since at least the time of *Coulomb* and his predecessors *Amontons*, *Desaguliers*, and *da Vinci*. Coulomb (1785) measured the time dependence of static friction, primarily of lubricated woods and also of metals, and derived a power-law aging relation

$$F = A + mT^n \quad [1]$$

where F is friction force, A represents cohesive or adhesive effects, T is time of contact, and m and n are empirical constants (see Dowson, 1979, for a summary). *Coulomb*'s value of n was ≈ 0.2, which for his range of available contact times is consistent with recent works in which aging is generally written in a log form: $F = A + m \ln(T)$. *Coulomb* refined his aging relation to account for short- and long-time saturation effects and suggested alternative forms to eqn [1]. In addition, he made a connection between aging and his second empirical friction law, which has been used widely as a brittle failure criteria in studies of faulting (see, e.g., Belardinelli et al., 1999; Baumberger and Caroli, 2006, and references therein):

$$F = A + \mu P \quad [2]$$

where P is normal force and μ is a coefficient of friction. *Coulomb* noted that the aging relation [1] provided an explanation for the difference between static friction and kinetic friction in eqn [2] and proposed a physical mechanism involving interpenetration of fibrous surfaces for time-dependent static friction.

The distinction between static friction and kinetic friction defines the simplest friction law. However, as *Coulomb* and many subsequent workers have noted, it involves a nonphysical singularity upon the initiation of sliding. Rabinowicz (1951, 1956, 1958) resolved this problem and inspired the next generation of friction laws by showing that friction evolves over a characteristic sliding distance D_c, which he related to physical characteristics of asperity contacts. *Rabinowicz*'s insights became formulated as a slip-weakening friction law (e.g., Ida, 1972; Palmer and Rice, 1973).

However, slip-weakening friction only complicates matters regarding frictional healing. That is, now in addition to asking how static friction is reset following failure, we may ask how the characteristic sliding distance is reset. Is D_c a property of the static or the dynamic contact, and is it therefore expected to vary with either time or sliding velocity? Moreover, once we admit time dependence of static friction and velocity dependence of kinetic friction, how, if at all, do these effects relate to resetting of static frictional strength following failure or to the onset or arrest of unstable slip?

The rate and state friction (RSF) laws were developed to address these issues (Dieterich, 1972, 1978, 1979; Rice and Ruina, 1983; Ruina, 1980, 1983). In the context of the RSF laws, time-dependent frictional healing and slip-dependent frictional healing are manifestations of the same processes that lead to velocity-dependent kinetic friction and memory effects. There are two distinct views of healing (see, e.g., Bayart et al., 2006; Beeler et al., 1994; Dieterich and Kilgore, 1994; Karner and Marone, 1998, 2001; Linker and Dieterich, 1992; Marone, 1998a; Ruina, 1983). In one, which we refer to as the *Dieterich law*, healing is fundamentally a time-dependent process involving growth of asperity contact area (e.g., Dieterich and Kilgore, 1994, 1996a,b) or increase in contact junction strength (Li et al., 2011; Tullis, 1996). In the alternative view, referred to as the *Ruina law or slip law*, any evolution of the physical surface, including healing, requires slip (e.g., Ruina, 1980, 1983) and the healing rate is zero for true stationary contact.

In both views of friction state evolution, frictional contacts have an average age determined by the inverse of sliding velocity divided by D_c, and macroscopic frictional strength increases with contact junction age. However, whereas time-dependent healing is apparently straightforward – if one imagines contact growth and strengthening via elastoplastic processes – the mechanisms that might cause slip-dependent healing are less evident. In the *Ruina law* interpretation, slip leads to both healing and weakening (as discussed more fully in the succeeding text). For multicontact solid surfaces, slip-dependent healing could arise from ploughing or shear-enhanced indentation, and in granular materials, it might arise from shear-enhanced consolidation. However, existing experimental evidence is limited and equivocal, with some results supporting the *Dieterich law* interpretation and others supporting a *Ruina law* interpretation (Bayart et al., 2006; Beeler et al., 1994; He et al., 1999; Karner and Marone, 1998; Li et al., 2011; Linker and Dieterich, 1992; Losert et al., 2000; Nakatani and Mochizuki, 1996; Richardson and Marone, 1999). We address this issue in the succeeding text with new data and analysis.

In previous studies of frictional healing, the effect of loading velocity was not widely appreciated (for exceptions, see Beeler et al., 1994; Johnson, 1981; Kato et al., 1992; Marone, 1998b). Measurements of frictional aging were focused primarily on time-dependent static friction, and they were carried out either in slide-hold-slide (SHS) tests or by bringing two surfaces in contact and waiting for a predetermined time before applying a shear load (e.g., Brockley and Davis, 1968; Coulomb, 1785; Dieterich, 1972). Details of the shearing rate were considered unimportant and not generally reported. However, experimental results show that aging and static friction vary with loading velocity, such that the static friction increases systematically with loading rate (Baumberger et al., 1999; Marone, 1998b). Moreover, in SHS tests, healing varies with shear stress (Berthoud et al., 1999; Karner and Marone, 1998, 2001; Nakatani and Mochizuki, 1996) and degree of granular consolidation (Bos and Spiers, 2002; Niemeijer and Spiers, 2006; Niemeijer et al., 2008, 2010; Richardson and Marone, 1999). Thus, measurements reported without such details are of limited value. Conversely, each of these variables can be usefully exploited as laboratory control parameters to study frictional healing, as was done by Beeler et al. (1994).

4.05.1.3 Repetitive Frictional Failure and the Upper Stability Transition from Stable to Unstable Sliding

A second prerequisite for repeated stick-slip failure (in addition to frictional healing) is that the fault exhibits rate-weakening friction – that is, friction decreases with increased sliding velocity. In its simplest form, the conditions for nucleation of unstable slip can be described by the force balance in a 1-D elastic system (e.g., Gu et al., 1984):

$$K < K_c = \frac{(b-a)\sigma'}{D_c} \quad [3]$$

where K is the effective elastic stiffness of the loading system, $(b-a)$ is the friction rate parameter, and σ' is effective normal stress. The RHS of eqn [3] defines the 'critical stiffness,' K_c, such that unstable slip is expected if $K < K_c$ (e.g., Gu et al., 1984; Ranjith and Rice, 1999).

A great deal of work has focused on defining the frictional rate dependence of natural fault materials and synthetic analogs, with the goal of testing hypotheses proposed to explain the upper transition from shallow aseismic creep to seismogenic, stick-slip behavior along active tectonic faults (e.g., Boulton et al., 2012; Carpenter et al., 2012; Collettini et al., 2009; den Hartog et al., 2012a,b; Ikari et al., 2007, 2009, 2011; Marone and Scholz, 1988; Moore and Saffer, 2001; Saffer and Marone, 2003). Because phyllosilicates are abundant in many fault zones (e.g., Deng and Underwood, 2001; Schleicher et al., 2006; Vrolijk, 1990; Vrolijk and Van der Pluijm, 1999), and commonly exhibit rate-strengthening behavior, these studies have been sufficient to rule out a range of hypothesized compositional controls on the upper stability transition (e.g., Saffer and Marone, 2003). Recent studies have shown that thermally activated deformation mechanisms and/or the onset of quartz deposition, associated with clay transformation and pressure solution, may drive a transition to velocity-weakening behavior at ~150 °C, coinciding approximately with the upper stability transition in subduction zones (e.g., den Hartog et al., 2013; Ikari et al, 2007).

The other elastic and frictional properties in eqn [3] have been less extensively explored but play an equally important role in governing frictional stability. In **Figure 3**, we provide one example illustrating these combined effects of friction rate dependence, the characteristic slip distance D_c, and loading system stiffness, K. We consider a subduction zone example, parameterized with laboratory-derived friction data from clay and clay–quartz mixtures analogous to natural fault gouge (Ikari et al., 2007), with the effective stiffness of the loading system defined by elastic constants measured for subduction zone sediments over a wide range of stresses (Gettemy and Tobin, 2003; Knuth et al., 2013). In this example, we predict a transition to unstable slip, based on the point at which the local stiffness K around a growing dislocation (e.g., Gu et al., 1984; Roy and Marone, 1996) drops below K_c. This occurs at a depth of ~4–6 km – comparable to the depth of the observed upper aseismic–seismic transition (cf. **Figure 1**) (e.g., Byrne et al., 1988; Hyndman et al., 1997; Moore and Saffer, 2001).

The results in **Figure 3** show that concomitant changes in elastic stiffness, rate weakening, and D_c associated with consolidation and lithification may play a central role in determining the updip limit of the seismogenic zone (Marone and Scholz, 1988; Saffer and Marone, 2003; Scholz, 1988a,b). Fundamentally, the processes within tectonic faults that drive changes in friction constitutive properties (e.g., rate dependence or D_c) with depth, such as lithification, cementation, mineral transformation, consolidation, or shear localization, are also likely to affect the loading system stiffness and fault healing behavior (cf. **Figure 3**) (e.g., Ikari et al., 2007; Moore and Saffer, 2001). Thus, three interlinked factors are necessary for effecting the transition from stable aseismic deformation, at shallow depths, to the seismic cycle of repeated frictional failure: (1) changes in frictional properties such that the critical stiffness K_c increases with depth in such a way that the criterion for instability is met (eqn [3]; **Figure 3**), (2) material stiffening brought about by consolidation of fault gouge and surrounding material, along with reduction of plastic strain, so that elastic stresses can be supported, and (3) the rate of frictional healing that must be sufficient to allow fault strength to increase at a rate that is higher than the rate of stress relaxation by creep and plastic deformation.

4.05.2 Laboratory Experiments

We introduce a limited set of new laboratory data collected in experiments designed to address frictional healing in a broad context. Direct shear experiments have played a key role in studies of fault zone friction. Here, we briefly describe this approach. The double-direct shear geometry (inset to **Figure 2**) is optimized for a biaxial load frame. This geometry eliminates the need for roller bearings, lubricated surfaces, or thrust bearings, as required for (single) direct shear (Dieterich, 1972; Hoskins et al., 1968). In the double-direct shear geometry, shear is accomplished using a three-block configuration in which a central bock is driven between two stationary side blocks (Dieterich, 1981; Ruina, 1983). The side blocks are shorter than the center block along the slip direction; thus, nominal contact area remains constant with slip.

4.05.2.1 Transducers, Control Electronics, and Data

The servo control system can be operated in either force or displacement feedback. Force measurements are typically made using custom-built, beryllium-copper load cells with active, full-bridge strain gauge networks for temperature compensation. The load cells and processing electronics are typically capable of resolving 1 N (a stress of 0.1 kPa on friction surfaces with nominal contact dimensions of 10 cm × 10 cm) over a range of forces up to 1 MN, and the servo control system is capable of maintaining stresses constant to this precision.

Displacements are measured with direct current displacement transducers (DCDTs) or linear variable differential transformers (LVDTs) and the primary control transducers are referenced to the loading rams, via fixed brackets (e.g., Karner and Marone, 1998). Thus, in displacement servo control, the position of a load point (the ram face) is controlled. Additional DCDTs are used to measure slip directly across the frictional surfaces in some cases. These are attached via brackets screwed directly to steel forcing blocks or via a spring-loaded clamp for granite. The clamp spans the length of the center block and thus average motion is measured in that case. When steel forcing

blocks are used, brackets are attached at a single screw point. Typically, the steel is assumed to be rigid in comparison with the gouge layer, which seems reasonable for quasistatic motions. Although DCDTs have unlimited resolution in theory, in practice, resolution is set by the maximum displacement range, processing electronics, and noise. These factors typically limit measurement precision to ~0.1 μm in most configurations.

Servo control amplifiers may be implemented in analog electronics or via computer control. Digital systems typically involve two analog-to-digital conversions: one to convert the signal from the analog transducer, to be used by the comparator and amplifier, and one to convert the digital output system to analog, to be used by the servo valve (for hydraulics) or stepper motor. Analog systems are typically faster but less flexible than digital systems. For complex load histories, and to control load point velocity, a digital signal and ramp generator may be used. Although the digital signal is typically converted to analog via a 16 or 24 bit converter, it is important that digital step size remains smaller than the resolution of interest. For friction experiments such as described here, load point velocities are implemented as digital ramps of 0.1 μm step size, and data are recorded with either 16 or 24 bit analog-to-digital converters and a nominal sampling interval of 10 kHz, averaged to recording rates ranging from 1 Hz to 1 kHz depending on loading velocity.

4.05.2.2 Materials

We focus on laboratory experimental examples using granular quartz available under the product name F-110 from the U.S. Silica. F-110 is a nearly pure quartz sand (99.8% SiO_2) with minor amounts of Fe_2O_3, Al_2O_3 (<0.1% each), and other oxides. Particles are rounded to subangular and the size distribution is as follows: 8% 53–75 μm, 25% 75–106 μm, 44% 106–150 μm, 18% 150–212 μm, and 4% 150–212 μm, with <1% beyond the upper and lower limits. For the experiments we describe here, gouge layers of uniform initial thickness were prepared with a leveling jig. The layers were hand-tamped to achieve uniform initial packing.

Gouge layers were sheared between two types of surfaces: Westerly granite with roughness produced by sandblasting and steel with corrugations machined perpendicular to the sliding direction. Granite surfaces had rms roughness of 80–100 μm (see Marone and Cox, 1994, for additional details on roughness characteristics for similar surfaces), and steel surfaces had a sawtooth roughness with amplitude of 0.9 mm and a spacing of 1 mm. In both cases, roughness was sufficient to inhibit slip at the boundary. The dimensions of the forcing blocks were as follows: side blocks 10 cm × 10 cm × 4 cm and center block 15 cm × 10 cm × 8 cm (Figure 4). Small, lubricated cooper shims were used beneath each side block to accommodate lateral motions. The center block has double thickness to account for the extra shear load it carries in the double-direct shear geometry and to ensure that strains are equal in each of the forcing blocks.

4.05.2.3 Velocity and Frictional Force Measurements

The equation governing position $x(t)$ of the central forcing block in the experimental configuration is

$$M\ddot{x} = K(V_{lp}t - x) - F \quad [4]$$

where M is the mass of the center block and gouge, K is the apparatus stiffness, V_{lp} is the load point velocity, and F is the frictional force. In a reference frame fixed to the side blocks, x is the fault position and $V = dx/dt$ is the fault slip velocity. For the conditions of the present experiments, inertia can be neglected and friction force is thus $F = K(x_l - x)$, where x_l and x represent the load point and fault position, respectively, relative to rest (e.g., inset to Figure 4). Since normal force N is held constant in the experiments, the nondimensional friction force can be written as

$$\frac{F}{N} = \frac{K}{N}(x_{lp} - x) \quad [5]$$

The frictional contact area is constant in the experiments; thus, $F/N = \tau/\sigma$ (where τ is shear stress and σ is normal stress), and it is convenient to plot the ratio τ/σ as a function of load point displacement (Figure 2). Because the conditions of present interest center around small perturbations from steady sliding, F is nearly always equal to μN, where μ is the coefficient of friction; thus, the force balance for perturbations to sliding reads

$$\frac{d(F/N)}{dt} = \frac{d\mu}{dt} = k(V_{lp} - V) \quad [6]$$

where k is the stiffness divided by normal stress (k equals 1×10^{-3} μm^{-1} for the experiments reported here). Parts of this development are contrary to simple notions about frictional strength, in which motion is zero until a threshold force is reached (e.g., static friction), but consistent with modern views based on the rate- and state-dependent laws in which macroscopic frictional force obeys a flow law that arises from microscopic thermally activated processes (e.g., Baumberger and Gauthier, 1996; Baumberger et al., 1999; Berthoud et al., 1999; Rice et al., 2001).

4.05.3 Laboratory Data

4.05.3.1 Experimental Boundary Conditions and Range of Major Parameters

Experiments are carried out at constant normal stress and a range of shearing rates under room temperature and humidity. Shearing rate was imposed as load point velocity and ranged from 1 to 300 μm s^{-1}. For the range of layer thicknesses (1–3 mm), the average shear strain rate ranged from 5×10^{-4} to 2×10^{-1} s^{-1}. At the initiation of each experiment, normal stress was raised to 25 MPa and maintained constant during shear (Figure 2). Dilation and compaction normal to the layers were measured continuously and these values are divided by two to obtain thickness variations for each layer.

Initial shearing resulted in significant compaction and strain hardening (Figure 2). Previous works have shown that the effect of strain hardening is reduced by cycling shear load to zero and back to failure (Marone et al., 1990; Shimamoto and Logan, 1981a), and this approach is used commonly. The results shown in Figure 2 are consistent with previous works on quartz, which show that frictional strength evolves systematically with accumulated shear strain due to comminution

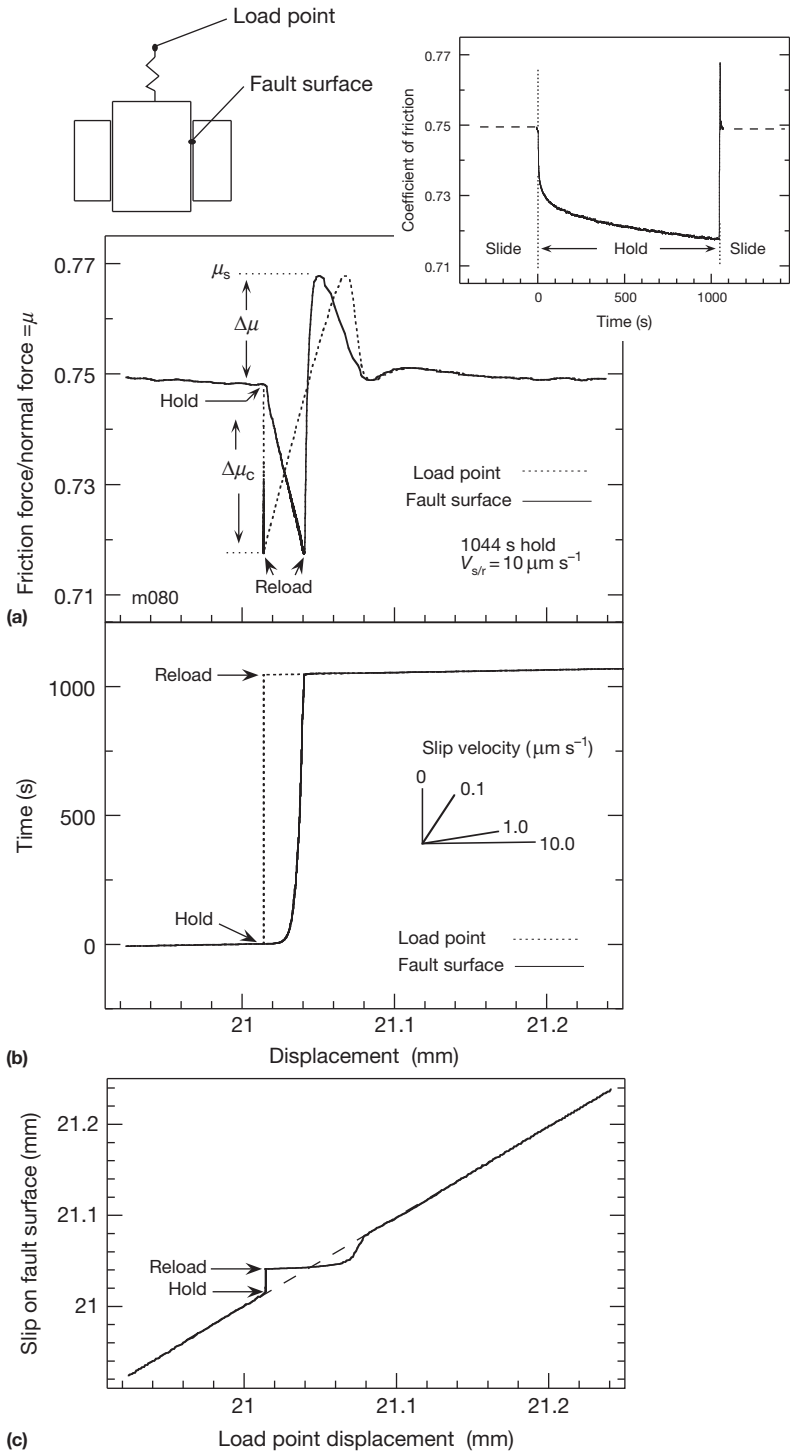


Figure 4 Details of a single SHS test. Upper left inset illustrates experimental geometry and displacement measurements. Upper right inset shows temporal history. V_l is zero during the hold and 10 μm s^{-1} before and after the hold. (a) Nondimensional friction force, $F/N = \mu$, is plotted versus load point displacement (dotted line) and slip measured at the fault surface (solid line). (b) Temporal history of the load point and fault position. Note that V reduces by over 2 orders of magnitude but remains finite during the hold period. (c) Relative motion between the load point and fault. From eqn [7], the two positions track one another when friction force is constant. $\triangle\mu$ and $\triangle\mu_c$ are the amount of frictional healing and creep relaxation, respectively.

and shear localization (Frye and Marone, 2002; Karner and Marone, 1998; Mair and Marone, 1999; Marone, 1998a; Marone and Scholz, 1989; Richardson and Marone, 1999; Saffer and Marone, 2003).

In any laboratory testing configuration involving direct or pure shear of confined layers, the layers undergo geometric thinning until shear becomes fully localized on boundary parallel shear bands (Rathbun and Marone, 2010, 2013; Scott et al., 1994). Initial compaction and quasilinear geometric thinning are clear in the data in **Figure 2**. Also, small amounts of quartz are lost from the front and rear edges of the layers if they are not jacketed or otherwise confined (Dieterich, 1981). Thus, the nominal contact area may decrease slightly with slip. Because of this, some investigators use lubricated plates at the front and back of the layers to eliminate loss along these edges (Mair and Marone, 1999). However, postexperiment inspection of the layers (e.g., Haines et al., 2009) indicates that gouge loss is limited to the outer edge of the perimeter (to a depth of about 1 layer thickness), and thus, the change in nominal contact dimension is small. Moreover, minor gouge loss occurs over large displacements relative to those of interest for frictional healing and second-order rate- and state-dependent friction effects. Accordingly, no artifacts of this effect have been reported.

4.05.3.2 SHS Tests, Creep, and Static Friction

SHS tests are commonly used to measure frictional healing and time-dependent aging under an applied shear stress (e.g., Beeler et al., 1994; Dieterich, 1972; Marone, 1998a). The tests measure stressed aging, as distinct from so-called unstressed aging, for which shear stress is zero during the waiting period (e.g., Baumberger et al., 1999; Berthoud and Baumberger, 1998; Berthoud et al., 1999; Bocquet et al., 1998; Géminard et al., 1999; Karner and Marone, 1998; Karner et al., 1997; Muhuri et al., 2003). In SHS tests, the hold period begins when V_l is set to zero and ends when loading resumes (**Figure 4**). Holds initiate during steady-state sliding, and thus, from eqn [6], $V = V_l$ at the initiation of a hold. We vary V_l, but keep the initial sliding velocity equal to the reload velocity. Thus, it is convenient to define a 'slide-reload' velocity $V_{s/r}$, which equals V_l both before and after a hold (e.g., Beeler et al., 1994).

Frictional creep begins immediately upon initiation of a hold, but because of finite apparatus stiffness, fault slip velocity does not reduce to zero. **Figure 4** shows detail of the creep and friction coefficient as a function of independent measurements of fault slip and load point displacement during a 1044 s hold. Note that during the hold period, sliding velocity V at the fault surface drops from 10 $\mu m\,s^{-1}$ to well below 0.1 $\mu m\,s^{-1}$ but remains finite (**Figure 4(b)**). During the hold period, friction decays logarithmically with time (upper inset to **Figure 4**). The parameter $\triangle\mu_c$ denotes the amount of creep relaxation during the hold. When reloading begins, fault slip is nearly zero initially, but slip initiates well before the peak in friction. This is consistent with previous works showing precursory slip prior to stick-slip failure (e.g., Scholz et al., 1972) and with the observation that fault slip rate does not reduce to zero (**Figure 4**). Our data indicate that frictional shear is continuous, albeit very slow, during hold periods conducted under shear load.

The peak friction value is generally taken to define the coefficient of static friction in a SHS test (Marone, 1998a). However, the fault surface is never truly static (e.g., Rabinowicz, 1956). From eqn [6], it is seen that peak friction merely represents the point at which fault slip velocity reattains its initial value (i.e., prior to the hold). Yet, for historical purposes and because it represents a critical yield strength, we measure peak friction and refer to it as the *static* friction.

Upon reloading, as slip continues after the friction peak, V overshoots V_l for a brief period (**Figure 4(c)**) due to finite apparatus stiffness. This velocity overshoot corresponds to the oscillation in friction after the peak (**Figure 4(a)**). It is important to note, however, that the maximum velocity during overshoot is only slightly above the loading velocity, 30–40 $\mu m\,s^{-1}$, compared to the loading velocity of 10 $\mu m\,s^{-1}$. Thus, the slight oscillation in friction force (which is also shown clearly in **Figure 5(a)** and subsequent figures) does not result in an inertial effect.

4.05.3.3 Frictional Healing, Creep Consolidation, and Dilation

Figure 5 shows data for two hold times and two loading velocities. Holds were carried out in pairs to assess measurement precision and experimental reproducibility. For a given velocity ($V_{s/r}$), the degree of both healing and relaxation increases with hold time, and for a given hold time, both $\triangle\mu$ and $\triangle\mu_c$ increase with velocity. These data illustrate a principal result of RSF. That is, frictional healing and relaxation vary systematically with both hold time and loading velocity (Kato et al., 1992; Beeler et al., 1994; Marone, 1998a,b).

Figure 5 also includes layer thickness measurements, shown as relative changes in porosity, where $\triangle\phi = \triangle h/h$ and h is layer thickness. Consolidation occurs during creep relaxation and dilation takes place when loading resumes (**Figure 5(c)**). Creep compaction mimics the time-dependent decay of friction, proceeding logarithmically in time (Beeler and Tullis, 1997; Marone, 1998b).

Upon reloading, dilation occurs in two distinct phases. An initial rapid dilation occurs over approximately the same displacement scale as attainment of the friction peak (**Figure 5**). This is followed by a second slower phase of dilation, which occurs as friction regains its steady-sliding value (**Figure 5(c)**). The duration of the second phase of dilation is longer for lower V_l, implying that it is driven by displacement rather than time (**Figure 5(c)**). For example, compare the 100 s holds at 3 and 1 $\mu m\,s^{-1}$ in **Figure 5(c)**. Like the friction data, consolidation and dilation associated with healing show a systematic relationship with both hold time and $V_{s/r}$. The initial dilation that occurs upon reloading influences frictional strength via the work necessary to expand the sample against the normal stress (e.g., Bishop, 1954; Edmond and Paterson, 1972; Marone, 1991; Marone et al., 1990), and recent works have accounted for this effect when evaluating frictional healing (e.g., Bos and Spiers, 2002; Frye and Marone, 2002; Karner and Marone, 2001; Niemeijer et al, 2008).

The effect of loading velocity on healing can be seen in a given experiment or by comparing separate experiments at different V_l (**Figure 6**). **Figure 6** shows data for a series of

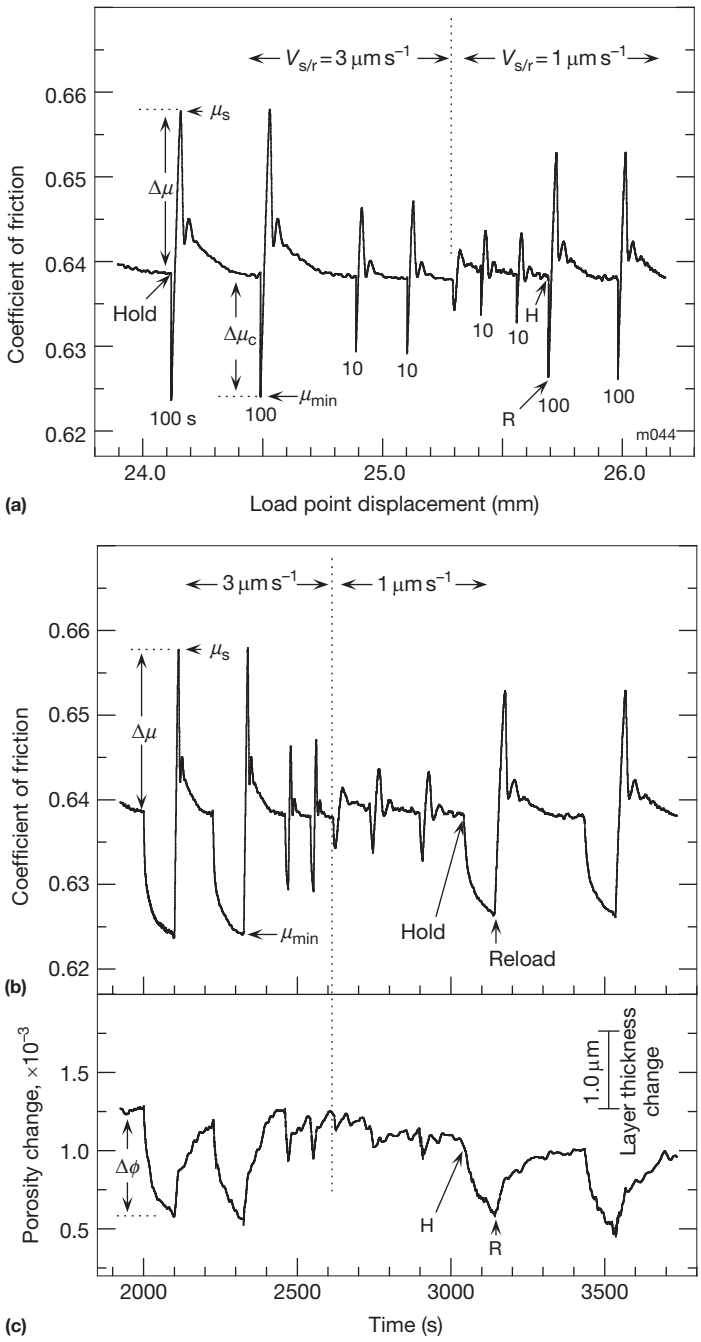


Figure 5 Healing characteristics for two loading velocities in a single experiment. (a) $\triangle\mu$ is the relative change in static friction and $\triangle\mu_c$ defines relaxation creep. Hold times are given below data. (b) Temporal characteristics of relaxation and reapproach to steady sliding. Note consistency of steady-state friction level for a given V_l. Friction exhibits velocity weakening (sliding friction is lower for higher V_l). (c) Gouge compacts during hold periods and dilates when loading resumes. Changes in layer thickness are shown as porosity change $\triangle\phi = \triangle h/h$, where h is layer thickness (scale bar gives $\triangle h$). Thickness data have been corrected for geometric thinning (e.g., see **Figure 2**) by removing a linear trend.

SHS tests from two separate experiments at different velocities. Precision of the experimental data and reproducibility are demonstrated. For a given hold time and velocity, healing and relaxation are nearly identical in repeat tests but differ consistently as a function of loading velocity. The degree of healing increases with V_l and, for a given hold time, creep relaxation is greater for higher V_l. The displacement required to regain steady slip also increases with hold time (e.g., **Figure 6(b)**).

Consistent frictional behavior is observed for a given displacement range; however, details of healing and relaxation evolve with net slip. For example, Richardson and Marone (1999) showed that for a given hold time and loading velocity, healing increases with net shear displacement, while creep

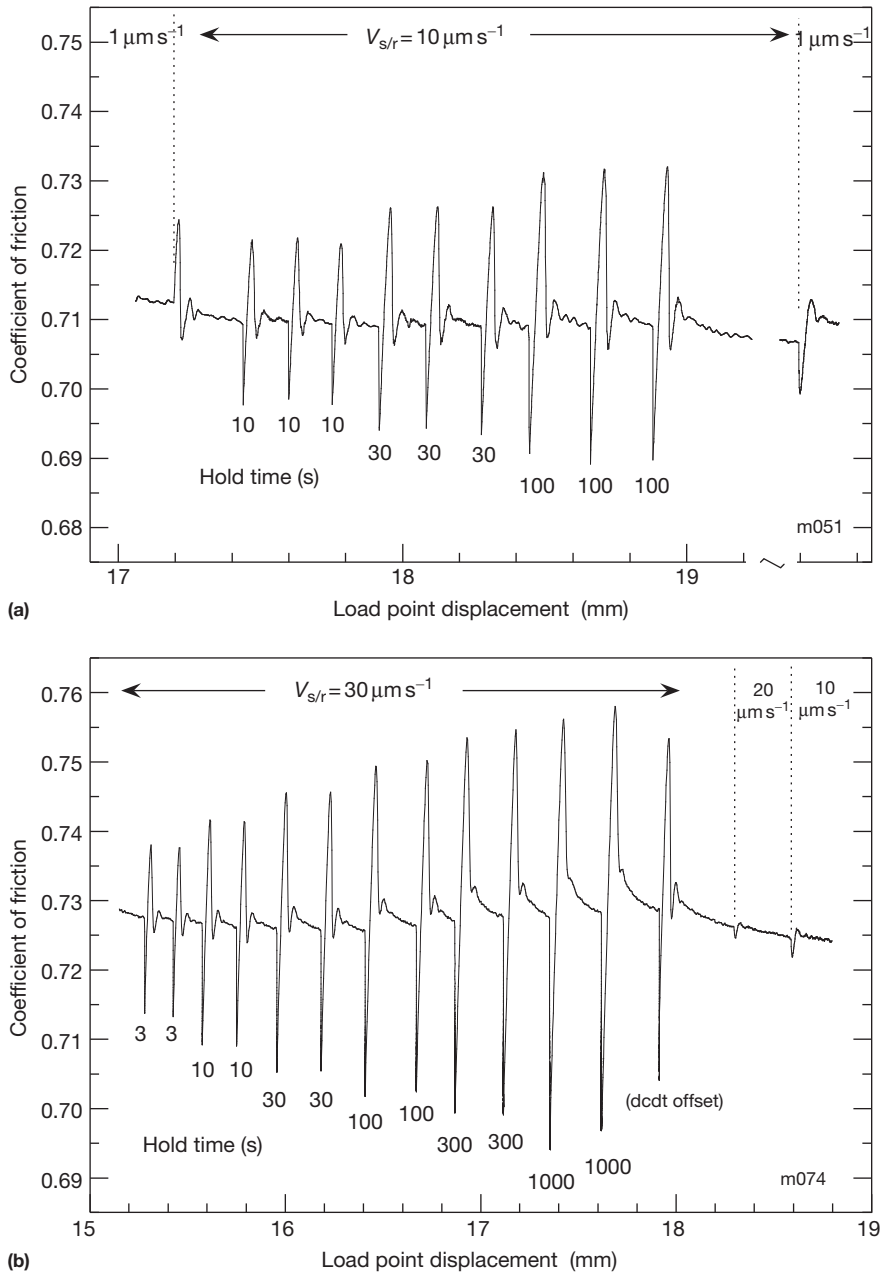


Figure 6 Healing characteristics for separate experiments and a range of hold times. (a) Three repeats at each hold time demonstrate experimental reproducibility. Healing and relaxation creep increase systematically with hold time. Note velocity-weakening frictional behavior as demonstrated by the response to a velocity step before and after the holds. (b) Data for loading at 30 μm s^{-1}. For a given hold time, healing and relaxation are systematically greater than for $V_{s/r} = 10$ μm s^{-1}. Note that the ordinate range is the same in both panels.

relaxation and compaction decrease. These trends are evident in the raw data (Figure 7). Figure 7 shows six sets of SHS with hold times ranging from 3 to 1000 s. Comparing the last pair of holds in each set, which are 1000 s, shows that $\triangle\mu$ is clearly larger and $\triangle\mu_c$ is smaller after a displacement of 30 mm than at 15 mm. The effect is also evident in the aging rate $\beta = \triangle\mu / \triangle\log_{10}(t_h)$, creep rate, and consolidation, which show systematic changes as a function of shear displacement (Figure 8). Figure 8 shows that for a 1000 s hold, healing increases by a factor of 1.5 as shear displacement increases from 13 to 28 mm. Existing data on granular quartz rich gouge show that healing parameters change nonlinearly with displacement, with most of the effect occurring prior to a shear of 20 mm (e.g., Richardson and Marone, 1999).

In an SHS test, creep relaxation occurs until strength meets or exceeds the applied stress. Thus, the displacement effect on healing and $\triangle\mu_c$ is consistent with the notion that the rate of time-dependent restrengthening (e.g., at contact junctions) increases as a function of net slip. This would result in less creep relaxation and greater static frictional strength.

Previous works included calibration experiments so that the effects of shear displacement could be removed (Richardson

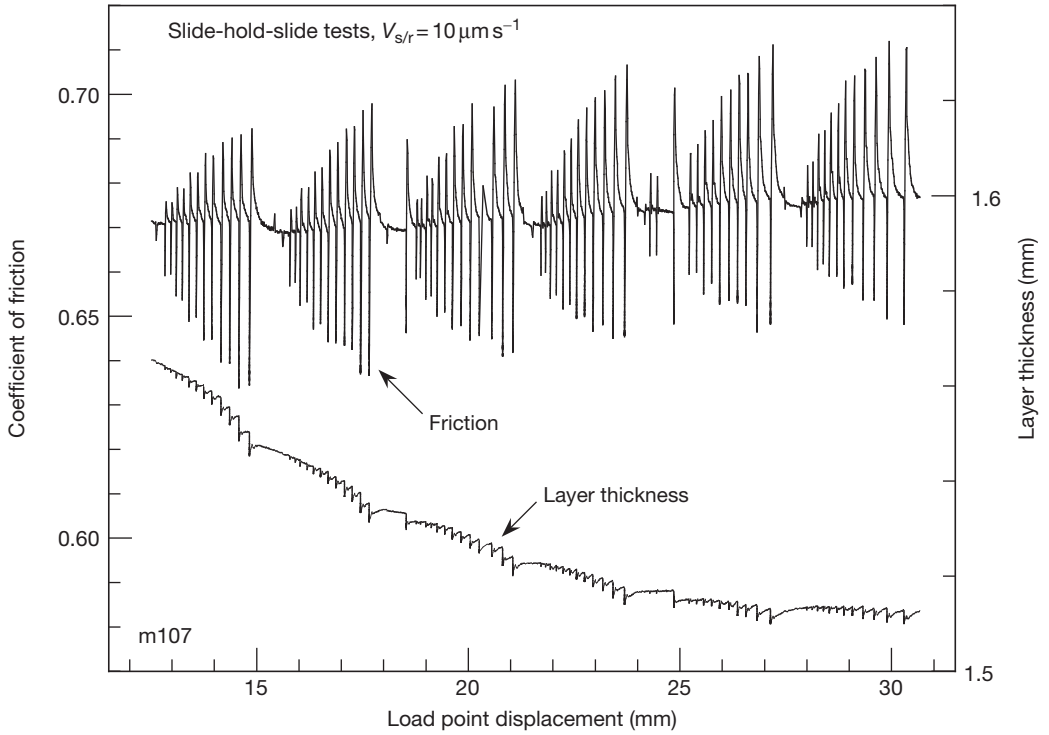


Figure 7 Six identical sets of healing tests showing the effect of cumulative displacement. The first set begins at a displacement of 13 mm and ends at 15 mm. In each set, holds of 3, 10, 30, 100, 300, and 1000 s are implemented in successive pairs. Note that for a given hold time, frictional relaxation and creep compaction decrease with slip, while healing and static friction increase with slip. The origin of the displacement effect is poorly understood but likely related to particle fracture and shear localization. Data from Richardson and Marone (1999).

and Marone, 1999). However, here, we consider data from only a narrow displacement range (e.g., Frye and Marone, 2002). Thus, in the detailed comparisons to follow, we use only data from net shear displacements of 15–19 mm.

4.05.3.4 Effect of Slip Velocity on Frictional Healing

Shearing velocity has an important effect on frictional healing. We explored a range of shear slip velocities to investigate the loading rate effect on healing. Figure 9 shows data from five experiments with loading rates from 1 to 100 $\mu m\,s^{-1}$. Within the scatter, healing increases linearly with the logarithm of hold time and velocity. The degree of healing increases by about the same amount for a tenfold increase in hold time as for a tenfold increase in loading velocity (Figure 9). Corresponding measurements of creep and compaction show similar trends to the healing data, with both increasing linearly with log hold time (Figure 10). Consolidation data are shown from two experiments, and others show similar trends. Creep during quasistatic holds tends to saturate for higher velocities and hold times. A clear V_l effect on $\Delta\mu_c$ is evident for holds of less than about 300 s (Figure 10). These data show the importance of controlling and reporting loading velocity when carrying out measurements of static friction and aging.

The effect of higher loading rates was investigated for V_l up to 300 $\mu m\,s^{-1}$. These data indicate that the velocity effect on healing extends to at least that level (Figure 11(a)). Data scatter is significantly lower for a single experiment than for multiple experiments (compare Figures 10 and 11), presumably due to minor experiment-to-experiment differences in load history or initial conditions, as suggested by previous work on solid granite surfaces and granite gouge (Beeler et al., 1994; Dieterich, 1972).

4.05.3.5 Creep, Precursory Slip, and Relocalization of Strain Following Healing

The slips that occur during creep relaxation and reloading are important aspects of frictional healing. Figure 12 shows measurements of these parameters. The inset shows details of an example slip history taken from the experiment shown in Figure 6(a) (the example is the second 100 s hold shown there). We define three quantities to characterize slip associated with healing: δ_h is the slip that occurs during a hold, from the beginning to point 1; δ_p is the precursory slip during reloading, from points 1 to 2; and δ_s is a measure of the residual slip necessary to regain steady-sliding friction, from points 1 to 3 (Figure 12). Creep slip δ_h increases roughly linearly with log hold time for a given velocity. These data mimic frictional creep (Figure 10(a)), which is not surprising given that the slip measurements are derived from measurements of load point displacement and friction force. Independent measurements of fault slip (e.g., Figure 4) show the same relations. For short hold times, δ_h increases with loading velocity. For holds greater than about 300 s, the velocity effect on δ_h is negligible (Figure 12(a)).

During reloading, precursory slip prior to peak friction ranges from 5 to 20 μm and decreases with increasing loading

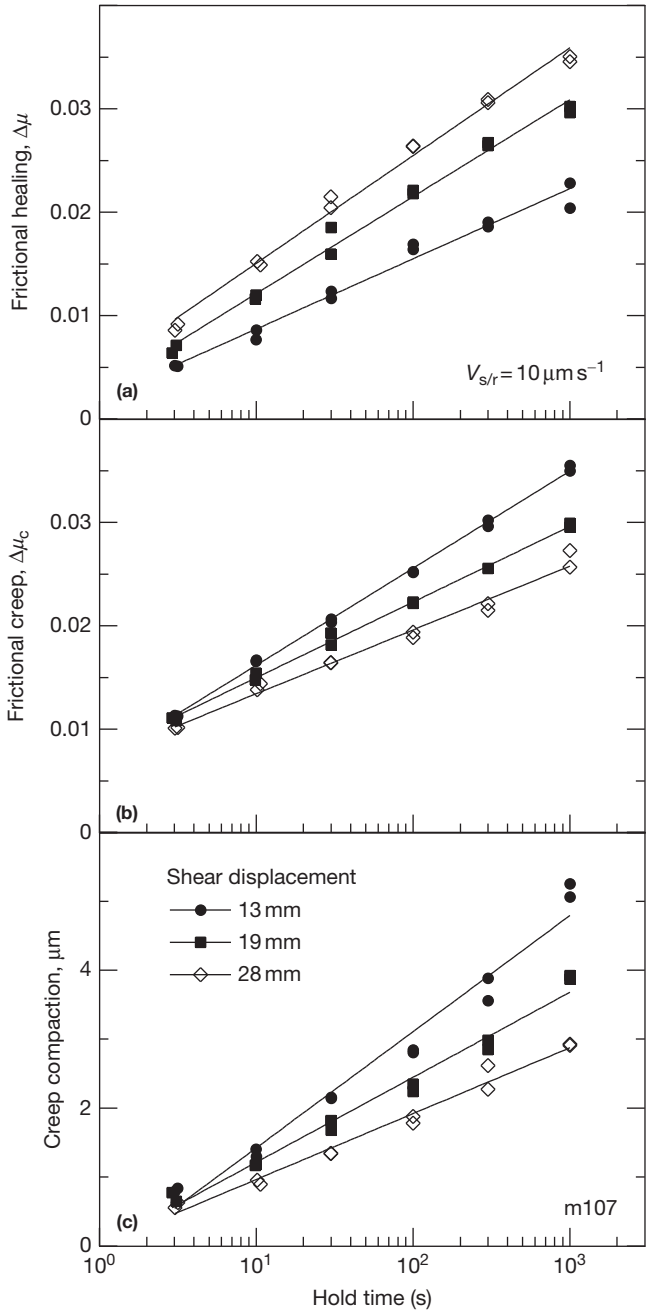


Figure 8 Systematic variations of (a) healing, (b) creep, and (c) compaction as a function of shear displacement. The first, third, and last sets of SHS tests from **Figure 7** are shown; labels correspond to the initiation displacement. For a given hold time, frictional relaxation and compaction decrease with displacement, while healing increases with displacement. Data from Richardson and Marone (1999).

velocity (**Figure 12(b)**). For shorter hold times, there is a significant velocity effect, such that for a given hold time, more precursory slip occurs at lower velocity. The effect is nonlinear as a function of log-t_h, and the reduction in δ_p saturates at about 5 μm before increasing. Clearly, additional data are needed to resolve this issue.

Healing also has longer-term effects on friction and reestablishment of steady sliding (e.g., **Figure 6**). The residual slip necessary to reestablish steady sliding seems to scale with hold time. We quantify these effects with the parameters δ_s. Ideally, residual slip would be defined as the slip necessary to regain steady-state friction, for example, the slip at which friction–displacement curves become flat again as in **Figure 6**. However, this is made difficult by minor variations in the base friction level as a function of accumulated slip and subtle differences in the postyield region of the friction–displacement curves. Thus, we chose to measure the slip at the first friction minimum following the peak (e.g., point 3 in **Figure 12** inset). Our measurements indicate that δ_s increases systematically as a function of hold time and loading velocity for shorter hold

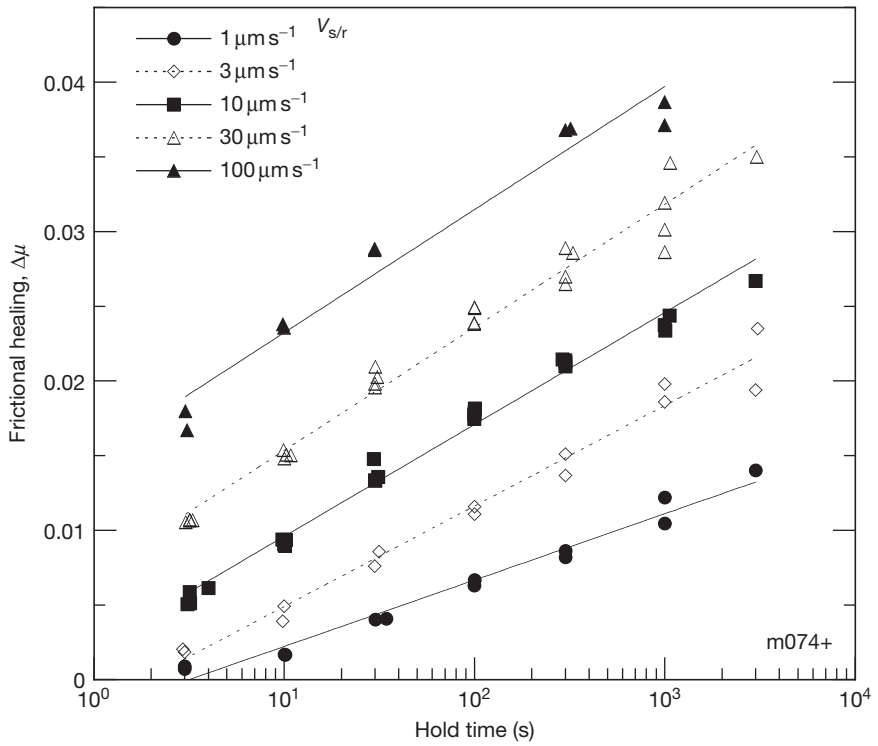


Figure 9 Data from multiple experiments showing the effect of loading velocity on frictional healing. Healing $\triangle\mu$ and, to a lesser extent, healing rate β increase systematically with loading velocity. The data indicate that 'static' friction and aging depend on both hold time and loading velocity and that the magnitude of the velocity effect is about equal to that of the effect of t_h; for example, a factor of 10 change in t_h produces about the same change in static friction as a factor of 10 increase in $V_{s/r}$.

times, below about 100 s, but that it tends to saturate for longer t_h. Part of this effect is likely related to the way we define of δ_s and perhaps other measures of residual slip would be useful.

4.05.4 Analysis and Discussion

The slip rate and state variable friction laws were originally developed to model velocity step tests (Dieterich, 1979; Ruina, 1983); however, the RSF laws can also model frictional healing and relaxation. Healing data illuminate different friction characteristics than velocity step tests, because healing generally involves large changes in slip rate and large perturbations from steady-state sliding. There are several advantages of modeling data from different types of tests, including the opportunity to evaluate the internal consistency of constitutive parameters and the RSF laws.

For the present discussion, we restrict attention to the RSF laws for constant normal stress and direct interested readers to other works that address the role of variable normal stress (Boettcher and Marone, 2004; Hong and Marone, 2005; Kilgore et al., 2012; Linker and Dieterich, 1992; Richardson and Marone, 1999). We evaluate two forms of the RSF laws. In each case, the coefficient of friction is written in terms of its dependence on slip velocity V and a state variable θ that represents physical characteristics of the shearing region (e.g., contact area in solid friction or porosity in granular media):

$$\mu(\theta_i, v) = \mu_o + a\ln\left(\frac{v}{v_o}\right) + b\ln\left(\frac{v_o\theta_i}{D_{ci}}\right) \quad [7]$$

where μ_o is a constant representing base-level friction at a reference velocity V_o, the parameters a and b are non-dimensional empirical constants, and D_c is a characteristic slip distance and the subscripts for θ and D_c represent separate evolution processes. Previous works have discussed generalization of eqn [7] for shearing within a region of finite thickness, as in the present experiments, rather than on a single surface (e.g., Marone and Kilgore, 1993; Marone et al., 2009; Sleep et al., 2000). In that case, V is replaced by a strain rate, and eqn [7] is written as

$$\mu = \mu_o + a\ln\left(\frac{\dot{\varepsilon}}{\dot{\varepsilon}_o}\right) + b\ln\left(\frac{\dot{\varepsilon}_o\theta}{\varepsilon_o}\right) \quad [8]$$

where $\dot{\varepsilon} = V_o/h_o$ (assuming spatially uniform strain), D_c is understood as a critical strain $\varepsilon_o = D_c/h_o$, and h_o is the total layer thickness. This form has the advantage of explicitly accounting for the effect of strain rate localization (see Rathbun and Marone, 2010; Sleep et al., 2000). That is, when strain rate is localized in a shear band of thickness h, the macroscopic and experimentally observed (apparent) value of the critical slip distance D_{app} is related to D_c for uniform strain by

$$D_{app} = \frac{hD_c}{h_o} \quad [9]$$

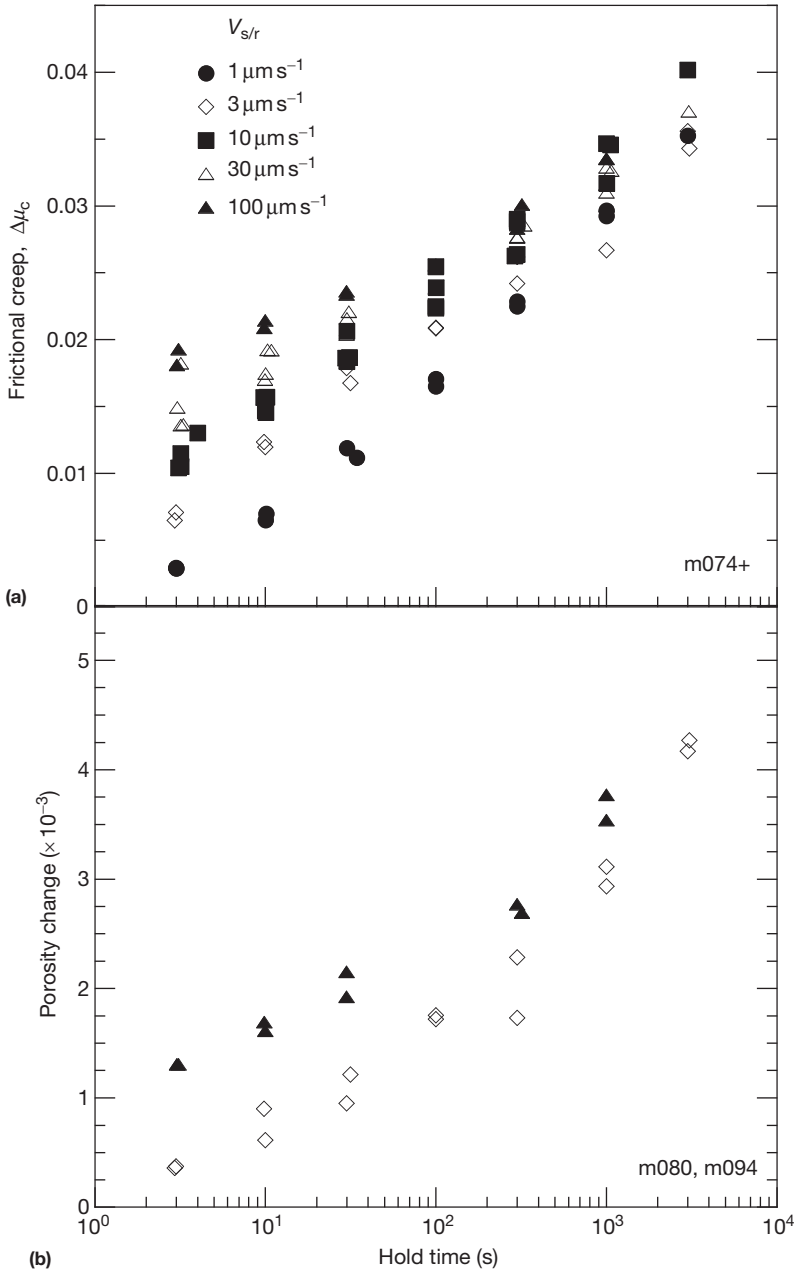


Figure 10 Details of the systematic effect of loading velocity on creep relaxation and compaction. (a) Creep increases with both hold time and loading velocity. The data indicate a tendency of creep relaxation to saturate at longer hold times and higher loading rates. (b) Compaction data are shown for two of the five loading velocities (others show similar trends). Consolidation during holds depends on both hold time and loading velocity in a way that mimics frictional creep.

as discussed by Ruina (1980) and shown experimentally by Marone and Kilgore (1993). To model strain rate localization and to illuminate the relationship between the RSF laws and other rheologies, eqn [8] can be rewritten as a flow law (King and Marone, 2012; Rice et al., 2001)

$$\dot{\varepsilon} = \dot{\varepsilon}_o \left[\frac{\varepsilon_o}{\dot{\varepsilon}_o \theta}\right]^{b/a} \exp\left[\frac{\mu - \mu_o}{a}\right] \qquad [10]$$

This formulation makes a clear connection between the RSF laws, thermally activated creep processes, strain localization, and strain delocalization during healing (e.g., Beeler, 2007; Berthoud et al., 1999; Heslot et al., 1994; King and Marone, 2012; Rice et al., 2001; Ruina, 1980; Sleep et al., 2000). For example, if μ is understood as a creep stress σ, then eqn [10] is readily rearranged as a logarithmic flow law $s \approx S \ln(\dot{\varepsilon}/\dot{\varepsilon}_o)$, where $S \approx a$ is a constant equal to $kT/\sigma_c\Omega$ (kT is energy in temperature units, σ_c is contact indentation strength, and Ω is activation volume for the thermally activated process), or a power-law creep relation $\sigma \approx A\dot{\varepsilon}^n$, if the exponent n is small, where A is a constant.

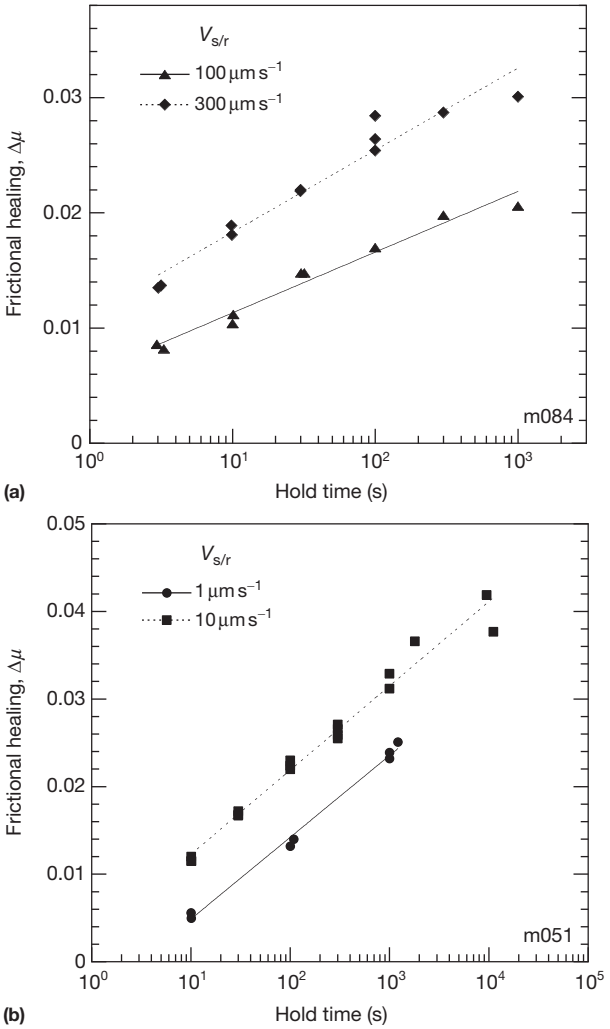


Figure 11 Healing data obtained from single experiments, for comparison of the data collections shown in Figure 9. (a) The velocity effect on healing extends to at least 300 μm s^{-1}. (b) Detailed comparison of two loading rates and hold times. The roughly log-linear relation between frictional healing and hold time seems to extend to 10^4 s.

Friction state evolution has been described by two forms:

$$\text{Dieterich law}: \quad \frac{d\theta}{dt} = 1 - \left(\frac{V\theta}{D_c}\right) \qquad [11]$$

$$\text{Ruina law}: \quad \frac{d\theta}{dt} = -\frac{V\theta}{D_c}\ln\left(\frac{V\theta}{D_c}\right) \qquad [12]$$

We focus here on the simpler forms of the RSF laws using macroscopic velocity, rather than strain rate. An important question is that of whether the first-order aspects of friction can be modeled with the RSF laws and, if so, whether the data distinguish between the two evolution laws [11] and [12]. There are two ways to address this question. One involves fitting the time series of a given SHS test (e.g., such as in Figure 5 or 6). The other (e.g., Beeler et al., 1994) is to fit the aggregate healing and relaxation data for a series of tests (e.g., such as given in Figures 10 and 11). We begin with the latter approach and address the former in the context of an analysis

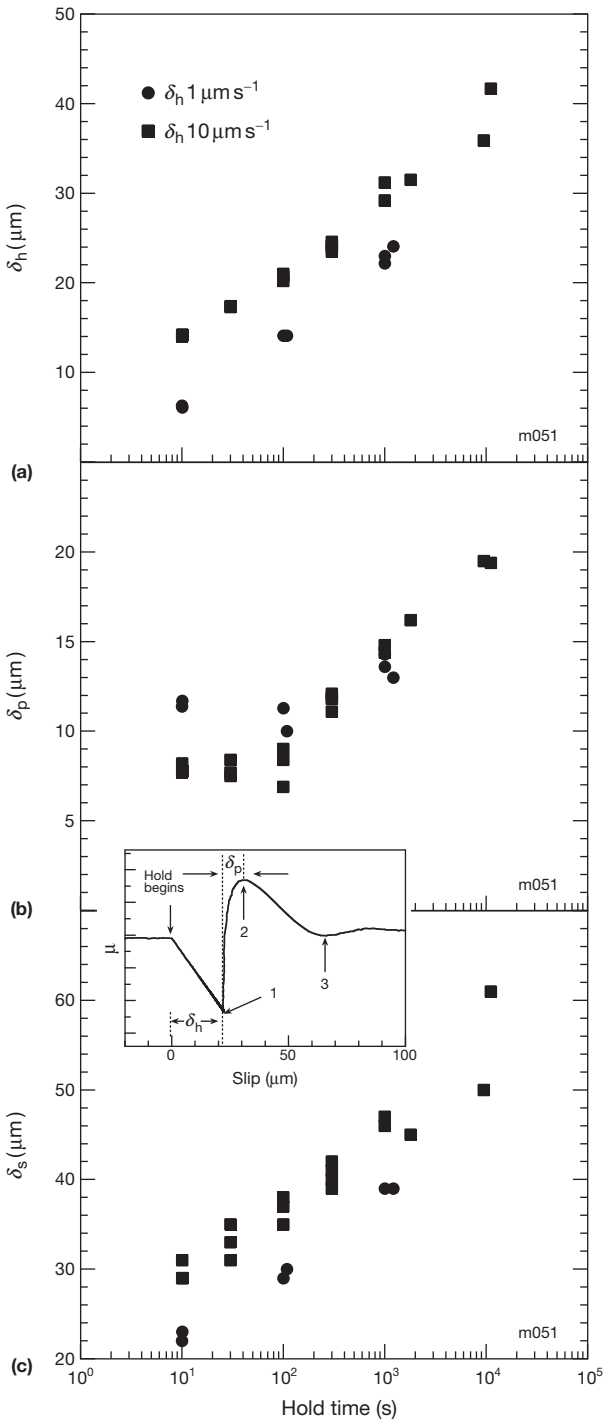


Figure 12 Details of slip characteristics during SHS tests. Inset shows representative friction-slip data set; hold begins at zero slip. δ_h is the slip that occurs during the hold, up to point 1. δ_p is the slip that occurs prior to peak friction, from points 1 to 2, and is referred to as precursory slip. δ_s is a measure of the slip necessary to reach the original value of sliding friction, from points 1 to 3. (A) The slip during δ_h increases systematically with log hold time and loading velocity. (B) Precursory slip decreases with hold time for low velocities. Note that δ_p is relatively independent of hold time and that it may be slightly lower for higher loading velocity. (C) δ_s increases nonlinearly with log hold time and, within the scatter, tends to be larger for higher loading velocity.

of the consistency of friction parameters obtained from different tests conducted in the same experiment.

4.05.4.1 Forward Modeling of Frictional Healing

Figure 13 shows numerical simulations of healing for each friction state evolution law. Forward models are calculated using typical values of the constitutive parameters for F-110 and the laboratory value of apparatus stiffness for the tests described earlier. The coupled equations [7] and [11] or [7] and [12] are solved using a fifth-order Runge–Kutta method. The inset provides details of the slip history for a single case, along with definitions of the healing parameters (these are the same as measured in the laboratory). The main panels in Figure 13 show healing and relaxation as a function of hold time for two velocities. The simulations mimic several aspects of the experimental data, including a velocity effect on healing and a region within which the healing parameter $\Delta\mu$ increases linearly with the log of hold time (Figure 13(a)). For a given hold time and loading rate, the *Dieterich law* predicts greater healing and lesser relaxation than the *Ruina law* (Figure 13). The *Dieterich law* also predicts that frictional creep relaxation saturates for long hold times. Note that the rate of change ($\Delta\mu_c/\Delta t_h$) drops significantly between about 10 and 100 s (Figure 13(b)). The simulations show a short-time cutoff effect at small hold times, which is consistent with experimental observations in the sense that, within measurement resolution, no healing is observed for hold times below about 1s for

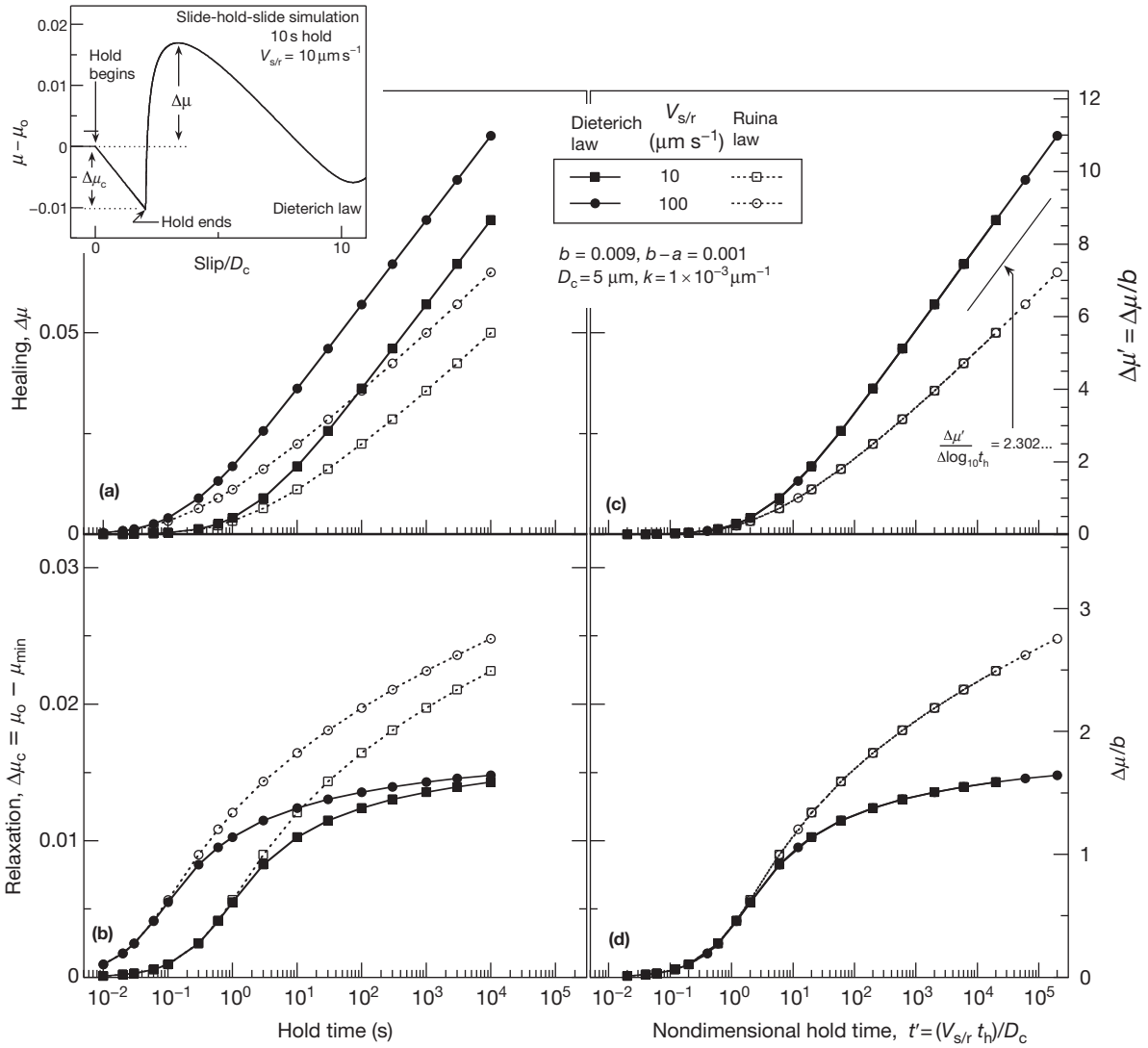


Figure 13 Numerical simulations of SHS tests showing healing and relaxation predicted by the RSF laws and elastic interaction. Inset to (a) shows details of a single simulation for which steady sliding is prescribed prior to a hold that begins at a normalized slip of 0 (the friction level is arbitrary and chosen as 0). The RSF laws predict that friction decreases during the hold due to creep and elastic interaction. Each point in the main panels shows results from a simulation. Four cases are shown corresponding to two velocities for each of the friction laws. Panels (a) and (b) show variations in the coefficient of friction. The same results are shown in nondimensional form in panels (c) and (d). Both laws show that healing and relaxation scale with loading rate and hold time, consistent with the experimental results. Simulations were carried out using the constitutive parameters given in the figure.

driving velocities of 10 μm s^{-1}. We note that data for the different velocities collapse when plotted versus nondimensional hold time t' (Figure 13).

The panels on the left side in Figure 13 show healing and relaxation data as changes in the coefficient of friction. The same simulation data are plotted in a normalized form in the right-hand panels and indicate that the distinction between time and velocity effects is accounted for when $V_{s/r}$ is normalized by an effective velocity $V_e = D_c/t_h$, to produce a nondimensional hold time t' (Figure 13). Both laws predict that, for a given set of constitutive parameters, healing (Figure 13(c)) and relaxation (Figure 13(d)) scale with the product of loading rate and hold time. The nondimensional form also shows that for the *Dieterich law*, the asymptotic healing rate β is equal to $b\ln(t_h)$ for long hold times (Figure 13(c)), as noted by previous workers (Beeler et al., 1994). This is expected, because for large t_h, V becomes very small and from eqn [11], $\dot{\theta} \approx 1$; thus, $\theta(t_h) \approx t_h$. This can be seen by taking $\mu = \mu_o$ during sliding prior to a hold; thus, healing is given by $\Delta\mu(t_h) = \mu_s(t_h) - \mu_o$, and substituting into [7] with $V_o = V_{s/r}$ (recall that peak friction is defined by the point at which $V = V_{s/r}$) yields $\Delta\mu(t_h) \approx b\ln(t_h)$.

4.05.4.2 Origin of the Loading Rate Effect on Healing

Numerical simulations were carried out to investigate the origin of the loading rate effect on healing. Figure 14 shows results of simulations carried out for a range of hold times (10^{-2} to 10^4 s) and slide/reload velocities $V_{s/r}$ (10–10^3 μm s^{-1}) using the parameters given in Figure 14(b). Figure 14(a) and 14(b) shows the complete time series computed for SHS tests, with data shown from initiation of the SHS to the peak value of friction. Figure 14(c) and 14(d) shows measurements of the healing slip parameters (see Figure 12) from a suite of simulations run using the values of t_h and $V_{s/r}$ shown.

In Figure 14(a) and 14(b), friction values are normalized by the state evolution parameter b and slip is normalized by D_c

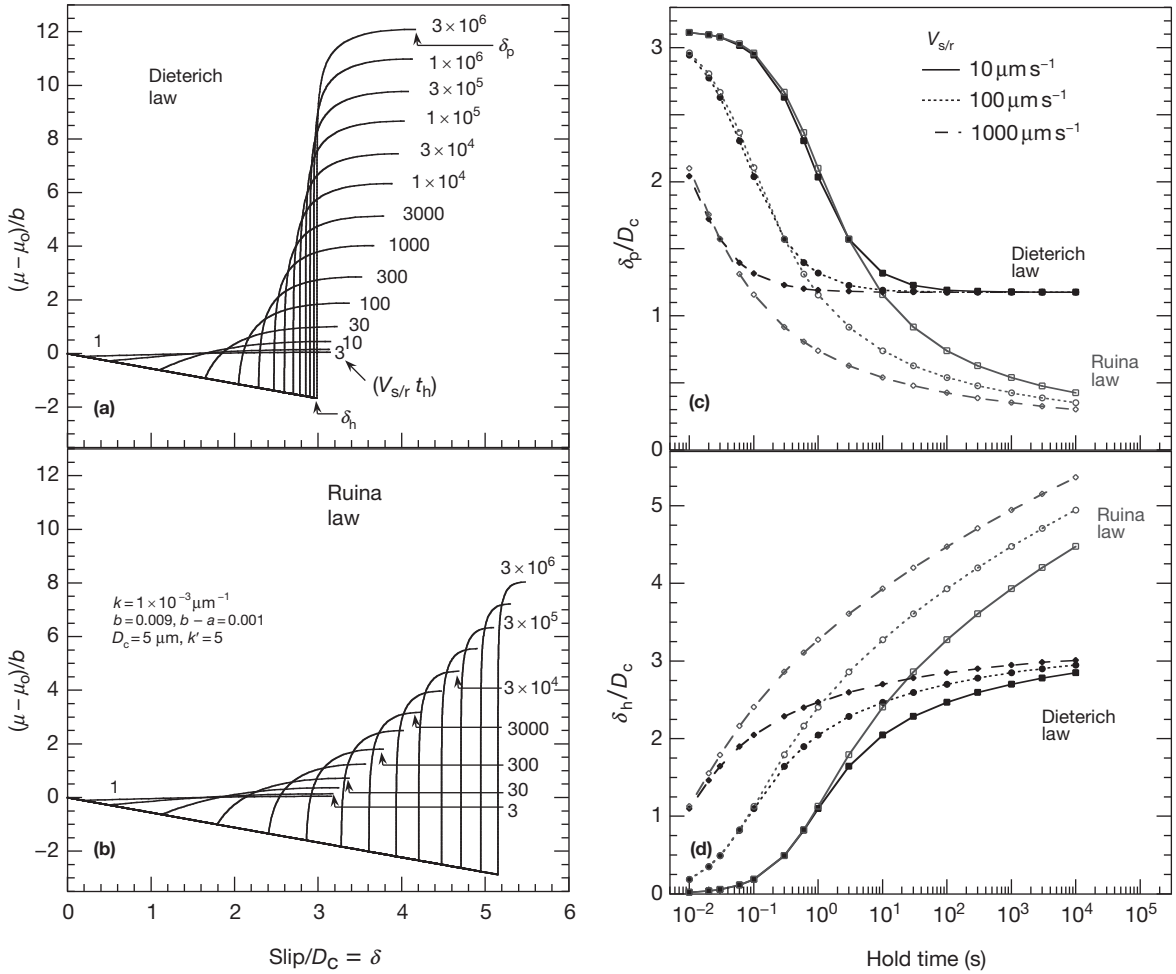


Figure 14 Characteristics of frictional slip during SHS simulations. Holds begin at zero slip and end when friction begins to increase (see inset to Figure 13(a)). The simulation is stopped when peak friction is reached. Panels (a) and (b) show results for the *Dieterich* and *Ruina* laws, respectively, plotted in nondimensional form. Curves are labeled by the product of loading rate and hold time and all simulations were carried out using the parameters given in panel (b). For a given velocity and hold time, the *Dieterich law* predicts less creep and slip during holds than the *Ruina law*. The parameters δ_h and δ_p (see Figure 12) indicate slip at the end of the hold and at peak friction, respectively. Panels (c) and (d) show the systematic changes in δ_p and δ_h as a function of velocity and hold time.

(Figure 14). Note that curves are labeled with the product of t_h and $V_{s/r}$. We also document the effect of loading rate and hold time on frictional creep and slip (Figure 14(c) and 14(d)). The slip parameters δ_h and δ_p are computed for a range of $V_{s/r}$ values and hold times.

The simulations indicate significant differences between the two state evolution laws (Figure 14). For a given hold time and loading velocity, the Dieterich law predicts significantly less creep slip and greater healing than does the Ruina law (Figure 14(a) and 14(b)). For the Ruina law, creep slip during the hold period exceeds $5D_c$ for the longest hold times. These simulations illuminate the origin of short-time, non-log-linear healing behavior (e.g., Figure 13(a)), in the sense that creep relaxation and healing are nearly zero for values of the product $t_h V_{s/r}$ below about 10.

We note that in an SHS test, creep occurs until strength meets or exceeds the applied stress. Thus, the prediction of greater creep slip for the Ruina law is consistent with the relation between slip and friction evolution (eqn [12]). That is, for the Ruina law, the rate of friction evolution goes to zero as slip velocity approaches zero. In contrast, the rate of friction state evolution becomes linear with time as V goes to zero for the Dieterich law, and thus, (1) creep relaxation is consistently smaller for the Dieterich law than for the Ruina law, and (2) creep relaxation becomes independent of loading velocity for the Dieterich law, for long hold times, but varies systematically with velocity for the Ruina law (Figure 14(d)).

The precursory slip during the reload following a hold period is quite similar for each law at short hold times. For hold times shorter than roughly 10 s, slip during the reload decreases with increasing hold time and slip velocity (Figure 14(c)). Smaller values of δ_p indicate a stiffer, more brittle response, which is consistent with the expectation that frictional contact junctions become stronger with increasing hold time. Similarly, the effect of $V_{s/r}$ on δ_p indicates a stiffer, more brittle response for higher loading velocities (Figure 14(c)). This is consistent with the friction direct effect, which produces a logarithmic increase in friction with velocity at constant state. For longer hold times and higher loading velocities, the Ruina law predicts a stiffer, more brittle response than the Dieterich law (Figure 14(c)). This is also evident in Figure 14(a) as the extended slip prior to reaching peak friction for the Dieterich law.

The effect of the $V_{s/r}$ on δ_p saturates for hold times longer than roughly 100 s for both laws, consistent with our laboratory results (Figure 12). For a given hold time and loading velocity, the Dieterich law predicts larger precursory slip than the Ruina law, consistent with the stronger dependence of the rate of state evolution on slip velocity (eqn [11]). Clearly, additional work is needed to better constrain these parameters.

Figure 15 shows a comparison of friction versus slip and the corresponding friction versus slip velocity (so-called phase plane plot: cf. Beeler et al., 1994; Rice, 1983) for two sets of simulations. We compare the responses for 100 s holds at two velocities for the Dieterich law and the Ruina law using the same friction parameters $a=0.008$, $b=0.009$, $D_c=5$ μm, and $k=1\times10^{-3}$ μm^{-1}. For the phase plane plots (Figure 15(b) and 15(d)), slip velocity is normalized by the initial velocity prior to the hold, so that $V_o=1$ or 10 μm s^{-1}. Note that friction decays to a lower velocity for $V_{s/r}=10$ μm s^{-1} than for 1 μm s^{-1} (Figure 15). This may seem counterintuitive: Why would slip velocity become lower if the initial velocity is higher? But this is an important part of the effect of velocity on healing rate. The rate of state evolution depends on slip velocity. Higher slip velocity results in greater slip, and together, they cause faster state evolution – that is, larger negative values. We note that differences in deceleration history during an SHS test, and during the seismic cycle, are part of the reason that frictional healing depends on loading velocity. This can also be seen by noting that for both laws, the change in the frictional state is larger for the 10 μm s^{-1} case than for the 1 μm s^{-1} case (Figure 15(b) and 15(d)).

The simulation results in Figure 15 show that friction increases approximately along a line of constant state upon reloading in an SHS test. Comparison of the friction curves with the fiducial lines of slope $a=0.01$ shows that friction begins to depart from a line of constant state at about two-thirds of the peak friction value for this range of stiffness and friction parameters.

For an SHS test, the RSF laws predict that peak friction occurs at the point at which slip velocity reaches $V_{s/r}$, as discussed earlier. This is clear in Figure 15(b) and 15(d). We note also that frictional overshoot, to velocities higher than the initial velocity, occurs for this range of parameters at both $V_{s/r}=1$ and 10 μm s^{-1}. This is seen by noting that V exceeds V_o after the peak in friction; then, V reaches a maximum value, followed by a minimum in friction and a local maximum at $V=V_{s/r}$ before friction and velocity drop back to their initial values (Figure 15). Note that in the recovery following overshoot, the friction curve increases at a higher rate than predicted for constant state. This is consistent with the combined effects of the direct effect and increased frictional state; friction increases faster than predicted by just the direct effect because of appreciable changes in state.

In addition to its role in defining the upper stability transition along tectonic faults (e.g., Figure 3), elastic stiffness also has an important effect on frictional healing and creep frictional relaxation (Figure 16). We show simulation results using values of the RSF parameters determined from velocity step tests, with some rounding to produce average values. Healing data are plotted versus a nondimensional stiffness given by k/K_c. Each data point in Figure 16 represents a full numerical simulation of an SHS test computed with a given hold time and elastic stiffness. Note that for $k<K_c$, frictional instability results, with V tending to large values (e.g., Ranjith and Rice, 1999). However, instability occurs after peak (static) friction is reached; thus, we obtain values of the healing parameters prior to numerical instability.

The simulation results show that frictional healing is larger for the Dieterich law (Figure 16), consistent with expectations from Figures 13 and 14. For values of $k\sim K_c$, healing increases with increasing stiffness for both laws (Figure 16(a)). For the Dieterich law, the stiffness effect on healing begins to saturate above values of about 50–100 K_c, whereas the Ruina law shows an optimum stiffness for healing with lesser healing for lower and higher stiffnesses (Figure 16(a)). For the Ruina law, maximum healing occurs at a stiffness of ~5 K_c, and healing tends toward zero for larger stiffnesses (Figure 16(a)). These results are consistent with the fact that the rate of state evolution goes to zero as V tends to zero for the Ruina law. For very large

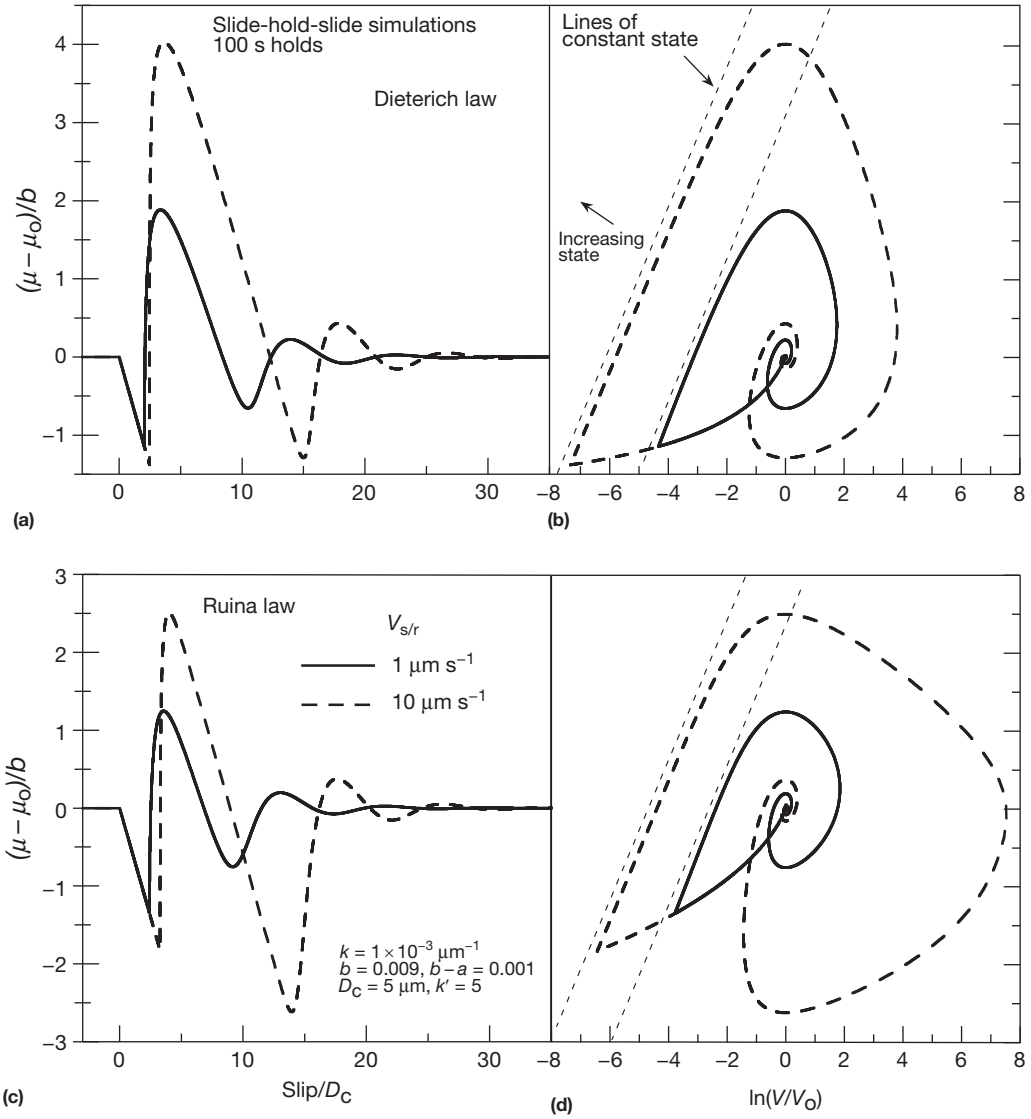


Figure 15 Numerical simulations of 100 s holds at two loading rates. Left-hand panels, (a) and (c), show friction versus displacement. Holds begin at slip = 0. Right-hand panels, (b) and (d), are phase plane plots showing corresponding data of friction versus slip velocity for the same holds shown in (a) and (c). In the phase plane plots, holds begin at $V = V_o$. After the hold, when friction has regained steady state, V returns to V_o. Note that the y-scale is the same in each row of plots. Simulations with the Dieterich law (a) and (b) show that loading rate has little affect on creep relaxation but that it has a pronounced affect on peak friction, whereas the Ruina law shows appreciably more slip and creep compaction at the higher initial loading rate. Dashed lines in the phase plane plots show lines of constant frictional state.

values of elastic stiffness, V drops to near zero very quickly for even a small frictional relaxation. This may be seen in Figure 16 (b), which shows frictional relaxation $\Delta\mu_c$ during the hold period. The values of $\Delta\mu_c$ are largest for the Ruina law and increase rapidly as stiffness increases (Figure 16(b)). This has important implications for natural faults, in the sense that the extent of healing predicted by the two friction laws is fundamentally different – and therefore would predict distinct interseismic healing behaviors as functions of loading system stiffness and therefore depth (e.g., Figure 3).

The influence of loading velocity on frictional healing is clear (e.g., Figure 9), but the origin of this effect, and the roles of the friction direct effect and state evolution, is complex (Figure 17). Each data point in Figure 17 represents a full numerical simulation of a 100 s SHS test computed with a given set of parameters. Six sets of parameters were used to explore the role of the friction parameters a, b, and $a-b$. We varied a at constant b, and b at constant a, with corresponding changes in $a-b$, and we also varied a and b so as to keep $a-b$ constant (Figure 17). The numerical results show that the origin of the velocity dependence of frictional healing differs for the two state evolution laws.

For the Ruina law, the velocity dependence of healing is due primarily to the direct effect (Figure 17(a)). For low values of a ($a = 8\times10^{-5}$), healing is essentially independent of slip velocity, and neither b nor $a-b$ has an appreciable effect on healing or creep relaxation (Figure 17). However, for a larger value of the direct effect ($a = 8\times10^{-3}$), both healing and creep

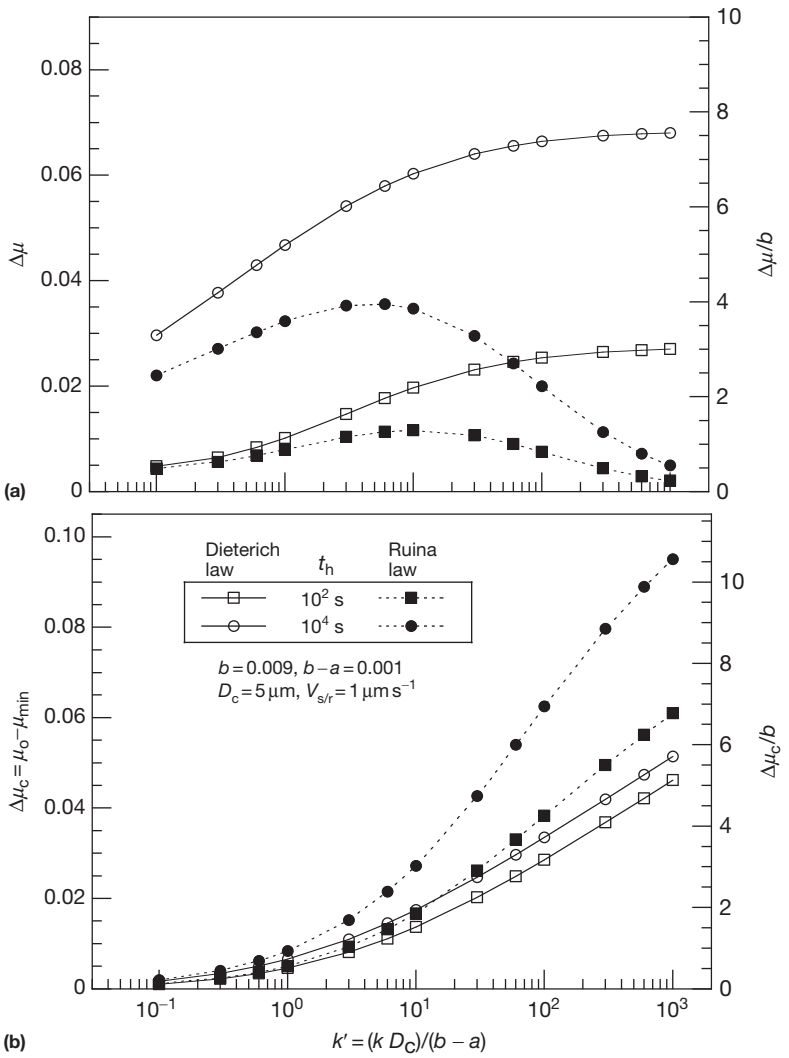


Figure 16 Numerical simulations showing the affect of stiffness on (a) healing and (b) creep relaxation for two hold times and each friction state evolution law. Each data point represents a full numerical simulation of a SHS test. All simulations were done using the RSF parameters given in panel (b). Stiffness is normalized by the critical stiffness k_c. For a given hold time and stiffness, frictional healing is larger for the Dieterich law and creep relaxation is larger for the Ruina law. For the Ruina law, healing reaches a maximum and decreases for stiffness greater than about $10k_c$.

relaxation increase significantly. For the range of velocities studied, frictional healing and the magnitude of creep relaxation vary directly with a. These results are consistent with the fact that the rate of state evolution tends to zero as V goes to zero for the Ruina law. Comparison of the friction phase plane plot for two cases with different values of the direct effect and constant b illuminates the role of the direct effect (Figure 18 (a)). The simulations in Figure 18 show that frictional creep is negligible for the lower value of the direct effect ($a = 8 \times 10^{-4}$); hence, slip velocity becomes very low and state does not change appreciably. In contrast, for a larger value of the direct effect ($a = 8 \times 10^{-3}$), creep relaxation is greater and thus slip velocity takes longer to drop, which means that state evolves faster and reaches higher values.

For these simulations, with $D_c = 5$ μm and initial slip velocity of 10 μm s^{-1}, the state term is initially 0.5 (Figure 18). For the case of $a = 8 \times 10^{-4}$, state is 1.04 at the end of the hold and 1.11 when peak friction is reached upon the reload. In contrast, for $a = 8 \times 10^{-3}$, state is 24.5 at the end of the hold and 6.1 when peak friction is reached upon the reload. Recall that state increases toward the upper left corner in a phase plane plot, but the rate of state change is scaled by the direct effect a (Figure 18). For the Ruina law, the direct effect influences frictional healing in an SHS test because creep relaxation scales directly with a and the rate of state evolution during the hold scales inversely with a. The velocity dependence of healing is produced largely by the direct effect a because larger values of a result in greater state evolution during the initial phase of deceleration, before velocity becomes negligible (Figure 18).

For the Dieterich law, both a and b influence the velocity dependence of healing (Figure 17). However, the effect of b is much greater than that of a and the signs of the effects differ. For the parameters in Figure 17, an increase of b of a factor of ~9 produces an increase in healing of nearly a factor of 10 at 1 mm s^{-1}. On the other hand, an increase of a by a factor of 100 results in a decrease in healing of only roughly 10% for the

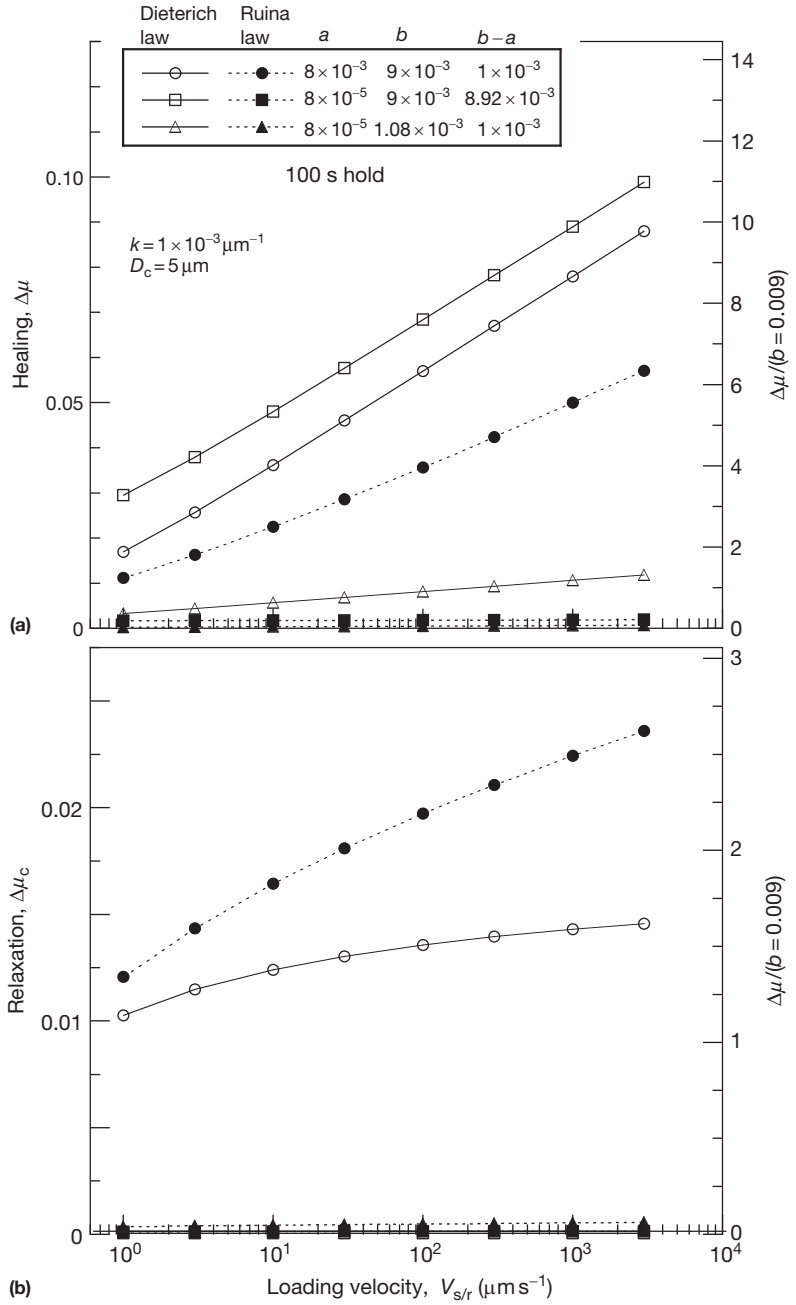


Figure 17 Numerical simulations showing the affect of RSF parameters and loading velocity on (a) healing and (b) creep relaxation for 100 s holds. Each data point represents a full numerical simulation of an SHS test. All simulations were done using the RSF parameters given in panel (a). Note that the three cases allow assessment of each RSF parameter: the friction direct effect a, the evolution effect, b, and the friction rate parameter $a-b$. For the Dieterich law, healing is dominated by the evolution parameter b and affected little by the direct effect a. For the Ruina law, healing is dominated by the direct effect a and affected little by the evolution effect b. For both laws, creep relaxation is affected primarily by a.

same velocity. These results are consistent with the fact that the rate of state evolution tends to 1.0 as V goes to zero for the Dieterich law. This means that the state term in eqn [7] continues to increase with hold time during an SHS test (Figure 17). As for the Ruina law, frictional creep scales directly with the friction direct effect for the Dieterich law (Figure 17(b)). However, because state continues to evolve as velocity goes to zero for the Dieterich law, state becomes large during the hold period and thus lower values of the direct effect, which cause a more rapid drop in slip velocity (Figure 18), produce larger frictional healing. For the simulations in Figure 18, state begins at a value of 0.5, and for the $a=8\times10^{-4}$ case, state is 100.4 at the end of the hold and 93.5 when peak friction is reached upon the reload. In contrast, for $a=8\times10^{-3}$, state is 85.5 at the end of the hold and 28.0 when peak friction is reached upon the reload (Figure 18).

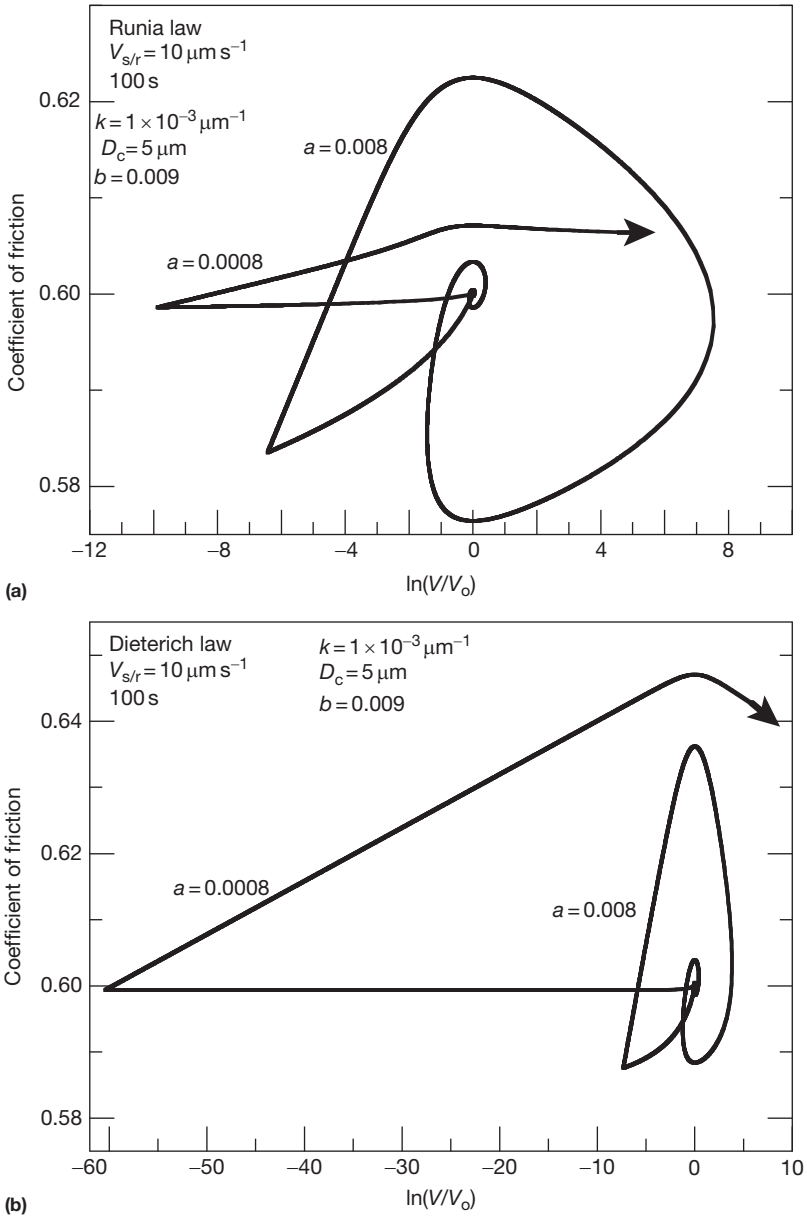


Figure 18 Numerical simulations using the (a) Ruina law and (b) Dieterich law to evaluate the affect of the friction direct effect on frictional healing and creep relaxation. See also **Figures 15** and **16**. Plots show the complete friction–velocity history for a 100 s SHS test. Comparison of the phase plane trajectories shows that frictional creep is negligible for the lower value of the direct effect ($a = 8 \times 10^{-4}$). Slip velocity during the creep relaxation stage of the SHS drops significantly in both cases, but the drop is particularly large for the Dieterich law where the velocity is $\sim 10^{-33}$ m s^{-1} at the end of the hold. The friction direct effect has a much larger influence on healing for the Ruina law than for the Dieterich law.

For the Dieterich law, healing varies inversely with the friction direct effect because larger values of a result in larger creep relaxation, which lowers the rate of deceleration and thus slows the rate of friction state evolution.

4.05.4.3 Recovery of Rate State Friction Parameters from SHS Tests

The RSF laws were originally developed from velocity step tests, and many previous works have documented the approach to recovering RSF parameters in these cases. However, RSF parameters can also be recovered from SHS tests (Figure 19). In this approach, the initial conditions for friction and state are taken from steady-state sliding prior to the beginning of the hold period. Then, frictional state and slip velocity are calculated during the hold and reload periods until sliding is once again at steady state. The numerical approach is essentially that of modeling two coupled velocity step tests: from $V_{s/r}$ to zero and back to $V_{s/r}$. The key difference is that the second 'velocity step,' which corresponds to the reload after the hold period, does not begin at steady state, but rather with values of state and velocity given by the hold.

Figure 19 shows detailed fits of the RSF laws to SHS time series. For clarity, we show model fits to data for a 10 s hold

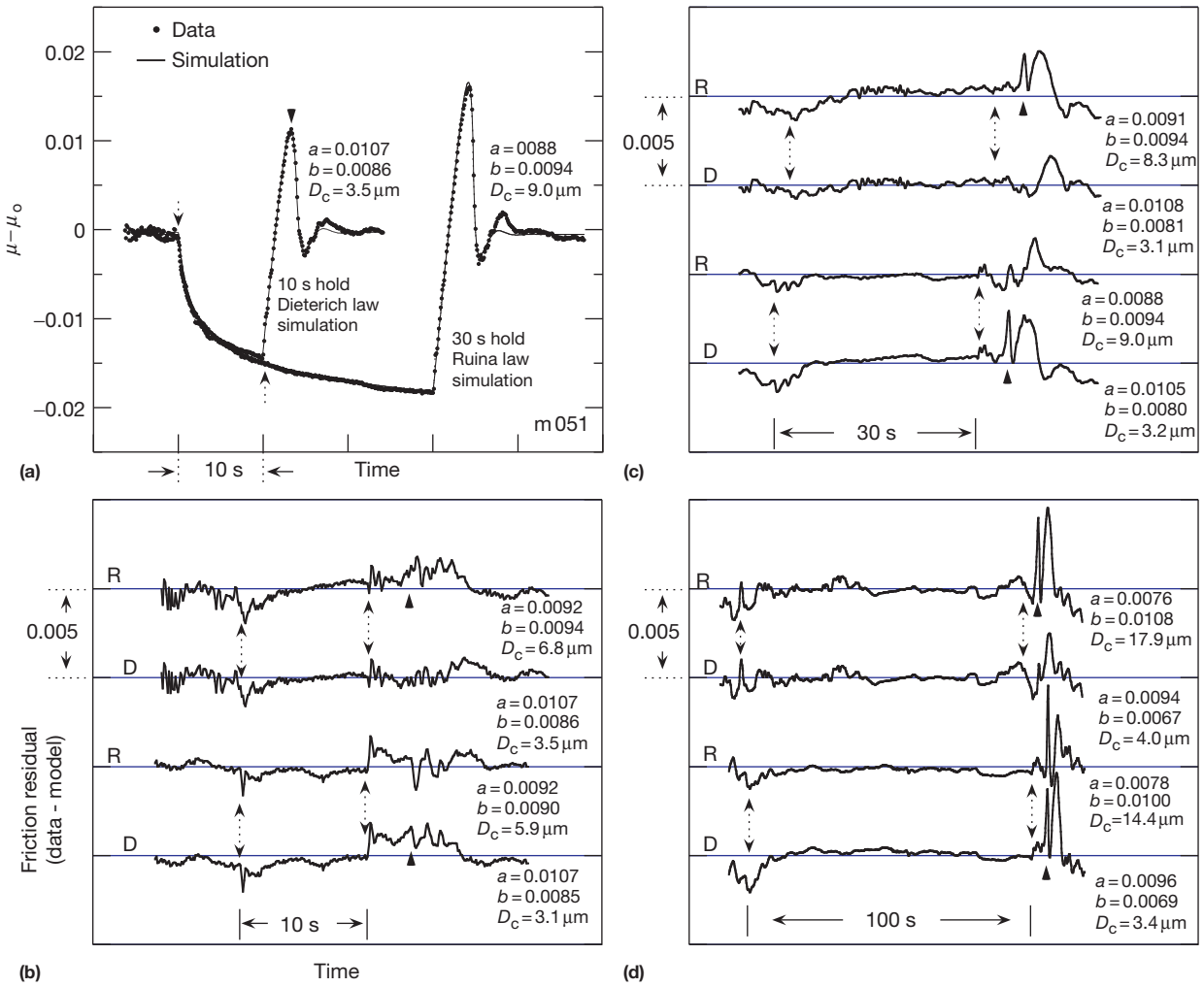


Figure 19 Friction data for SHS tests and best-fit numerical simulations using the RSF laws. (a) The change in friction from the steady state is shown versus time for two SHS tests. The 10 s hold is fit with the Dieterich law and the 30 s hold is fit with the Ruina law. RSF parameters are given for each case. Note that the RSF simulations fit the data quite well throughout nearly the entire SHS. Panels (b)–(d) show residuals computed by finding a least-squares best-fit RSF simulation to an SHS test, as in (a), and then subtracting the fit from the friction data. Each panel shows data for two SHS tests and the fits for the Ruina law and the Dieterich law. Arrows denote the beginning and end of the hold and the peak in friction – see panel (a) for corresponding arrows in the time series for the 10 s SHS. Note that the best-fit RSF parameters differ somewhat as a function of hold time (compare values in panels (b)–(d)) but that they are relatively consistent. In each case, the best-fit RSF parameters a and D_c are larger for the Dieterich law compared to the Ruina law, and the evolution parameter b is larger for the Ruina law.

and a 30 s hold together and offset the data so that the hold begins at the same time in each case (Figure 19(a)). The arrows denote the beginning and end of the hold period for the 10 s hold. Note that the friction creep curves, from the initiation of the SHS test, are identical (Figure 19(a)) and that each data set is reasonably well fit by the RSF simulation. We also show fits to six additional SHS tests, two for 10, 30, and 100 s holds, respectively (Figure 19). In each case, the same SHS data were fit with each state evolution law and we show the residual (data – model) as a function of time. Figure 19(b) shows data for two separate 10 s SHS tests, and the best-fit RSF parameters are quite similar between the data, for a given evolution law. Note that $a = 0.0092$ for both of the Ruina law fits (Figure 19(b)). The best-fit RSF parameters vary somewhat between the data for 10, 30, and 100 s, but they are reasonably close.

We may use the best-fit RSF parameters recovered from individual SHS tests to assess the extent to which they predict the effect of loading rate on frictional healing and creep relaxation. Figure 20 shows the laboratory data from Figure 9 plotted versus a normalized hold time formed by the product of the slip velocity $V_{s/r}$ and hold time t_h. The x-axis in Figure 20 has units of displacement, which we could normalize by D_c, but we chose not to because we do not know D_c a priori. The loading rate effect on healing and creep relaxation is quite well described by the normalized hold time. We also show two curves in Figure 20. These are predictions of the healing behavior computed from the RSF parameters obtained from fits to individual SHS tests (as in Figure 19). The predicted curves are computed from averaged values (Figure 19), and thus, they are expected to fit the healing (Figure 20(a)) and creep relaxation

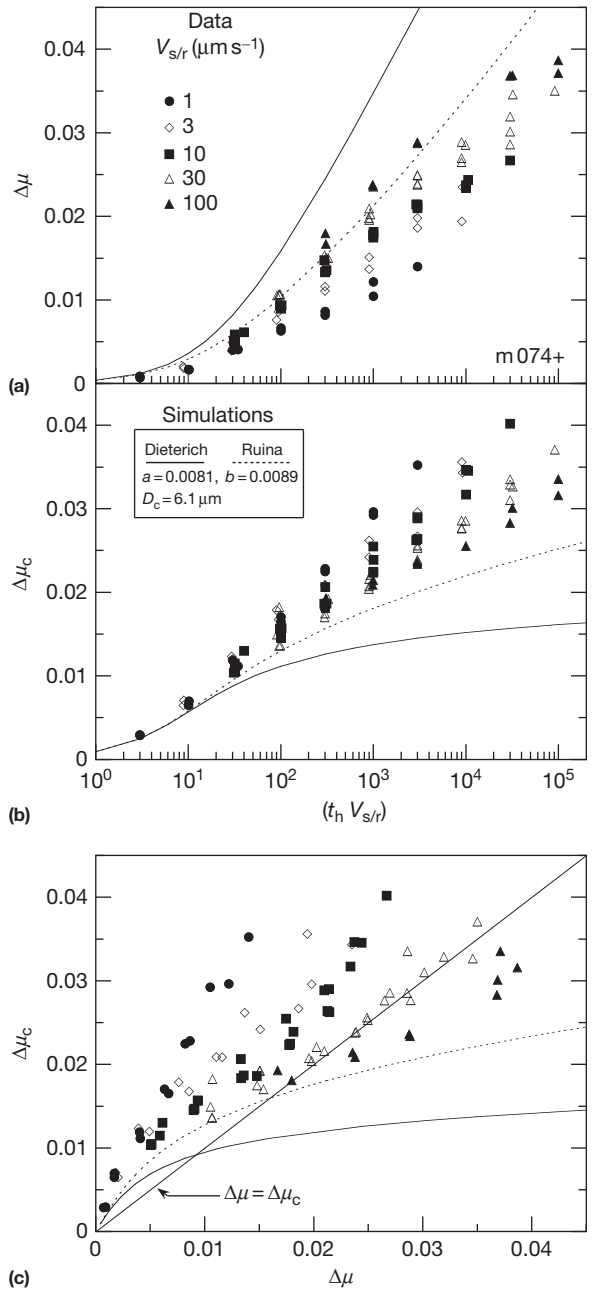


Figure 20 Friction data for SHS tests are plotted versus (a and b) the product of hold time and the slide-reload velocity $V_{s/r}$. Panel (c) shows healing data plotted versus creep relaxation. For each panel, lines show RSF predictions based on the parameters given in (b) obtained by fits to the individual SHS test (see **Figure 19**). Note that the effect of loading rate $V_{s/r}$ on healing is accounted for by the effective healing displacement ($t_h V_{s/r}$). The Ruina law prediction is closer to the trend of the laboratory data for each plot.

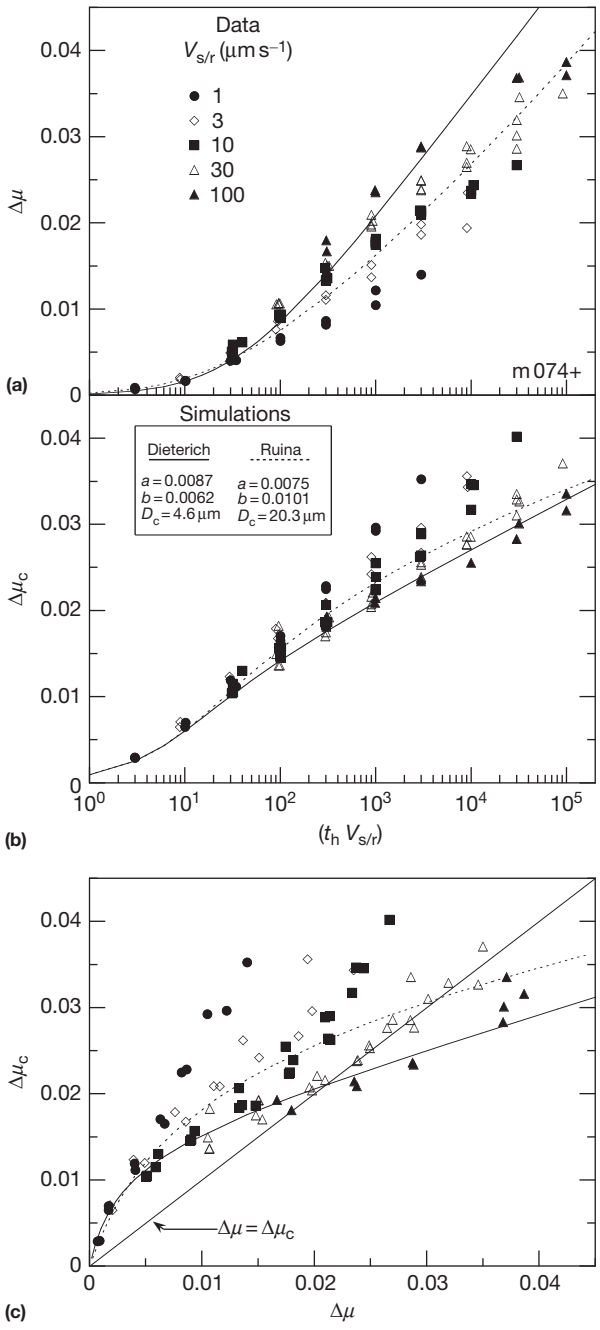


Figure 21 Same data as **Figure 20**. Curves show best-fit RSF predictions for each law. Note that the RSF parameters are similar to those in **Figure 20**, but these curves do a much better job of fitting the overall trend of the data. The Ruina law prediction is closer to the trend of the laboratory data for each plot.

(**Figure 20(b)**) data, at least approximately. While neither curve matches the data perfectly, the Ruina law prediction is much closer to the trend given by the laboratory data for both healing (**Figure 20(a)**) and creep relaxation (**Figure 20(b)**). **Figure 20(c)** shows data and model fits for creep relaxation versus healing. This representation shows that creep relaxation is larger than healing for small hold times but that creep tends to saturate for longer times, while healing continues to increase. This trend is captured by simulations with both state evolution laws, but the Ruina law is closer to the data trend.

One can also find best-fit RSF parameters for the complete SHS data set, at all loading rates, and we show such fits in **Figure 21**. The constitutive parameter values are not too

different from those in Figure 20, which is consistent with the trends shown in Figure 20.

4.05.4.4 Microphysical Explanation of Frictional Healing and Rate Dependence

If we assume that shearing in granular quartz occurs primarily by granular flow mechanisms, then it is reasonable that steady-state porosity is higher at higher sliding velocities, which would result in more compaction and relaxation during holds. The observations of greater healing, and, to a lesser extent, increased healing rates, could then be attributed to the larger magnitude of dilation upon reshear. Moreover, the observed changes in healing, compaction, and relaxation with increasing shear displacement (Figures 7 and 8; Richardson and Marone, 1999) are consistent with shear-driven comminution, grain size reduction, and shear localization.

4.05.5 Conclusions

Elastic stiffness has an important effect on frictional healing and creep frictional relaxation (Figure 16). We show simulation results using values of the RSF parameters determined from velocity step tests, with some rounding to produce average values. Healing data are plotted versus a nondimensional stiffness given by k/K_c. Each data point in Figure 16 represents a full numerical simulation of an SHS test computed with a given hold time and elastic stiffness. Note that for $k < K_c$, frictional instability results with velocity tending toward large values.

Fundamentally, the processes within tectonic faults that drive changes in friction constitutive properties (e.g., rate dependence or D_c) with depth, such as lithification, cementation, mineral transformation, consolidation, or shear localization, are also likely to affect the loading system stiffness and fault healing behavior (cf., Figure 3) (e.g., Ikari et al., 2007; Moore and Saffer, 2001). Thus, three interlinked factors are necessary for effecting the transition from stable aseismic deformation, at shallow depths, to the seismic cycle of repeated frictional failure: (1) changes in frictional properties such that the critical stiffness K_c increases with depth in such a way that the criterion for instability is met (eqn [3]; Figure 3), (2) consolidation of fault gouge and surrounding material so that elastic stresses can be supported, and (3) the rate of frictional healing that must be sufficient to allow fault strength to increase at a rate that is higher than the rate of stress relaxation by creep and plastic deformation.

Acknowledgments

We thank N. Beeler, B. Carpenter, M. Ikari, A. Niemeijer, and M. Scuderi for discussions and A. Niemeijer and N. Beeler for thorough, thoughtful reviews that improved the manuscript. This work was supported by NSF grants OCE-0648331, EAR-0746192, and EAR-0950517 and IGPP/LANL grant 162505-1.

References

Angevine CL, Turcotte DL, and Furnish MD (1982) Pressure solution lithification as a mechanism for the stick-slip behavior of faults. *Tectonics* 1: 151–160.

Baumberger T and Gauthier L (1996) Creep like relaxation at the interface between rough solids under shear. *Journal de Physique I France* 6: 1021–1030.

Baumberger T, Heslot F, and Perrin B (1994) Crossover from creep to inertial motion in friction dynamics. *Nature* 367: 544–546.

Baumberger T, Berthoud P, and Caroli C (1999) Physical analysis of the state- and rate-dependent friction laws: II. Dynamic friction. *Physical Review B* 60: 3928–3939.

Baumberger T and Caroli C (2006) Solid friction from stick slip down to pinning and aging. *Advances in Physics* 55(3): 279–348. http://dx.doi.org/10.1080/00018730600732186.

Bayart E, Rubin A, and Marone C (2006) Evolution of fault friction following large velocity jumps. *EOS, Transactions, American Geophysical Union*, F 2006.

Beeler NM (2007) Laboratory-observed faulting in intrinsically and apparently weak materials: Strength, seismic coupling, dilatancy, and pore fluid pressure. In: Dixon TH and Moore JC (eds.) *The Seismogenic Zone of Subduction Thrust Faults*, p. 692. New York: Columbia University Press, ISBN: 978-0-231-13866-6.

Beeler NM and Tullis TE (1997) The roles of time and displacement in velocity dependent volumetric strain of fault zones. *Journal of Geophysical Research* 102(22): 595–22609.

Beeler NM, Tullis TE, and Weeks JD (1996) Frictional behavior of large displacement experimental faults. *Journal of Geophysical Research* 101: 8697–8715.

Beeler NM, Tullis TE, and Weeks JD (1994) The roles of time and displacement in the evolution effect in rock friction. *Geophysical Research Letters* 21: 1987–1990.

Beeler NM, Tullis TE, and Goldsby DL (2008) Constitutive relationships and physical basis of fault strength due to flash heating. *Journal of Geophysical Research* 113. http://dx.doi.org/10.1029/2007JB004988.

Behnsen J and Faulkner DR (2012) The effect of mineralogy and effective normal stress on frictional strength of sheet silicates. *Journal of Structural Geology* 42: 49–61. http://dx.doi.org/10.1016/j.jsg.2012.06.015.

Belardinelli ME, Cocco M, Coutant O, and Cotton R (1999) Redistribution of dynamic stress during coseismic ruptures; evidence for fault interaction and earthquake triggering. *Journal of Geophysical Research* 104: 14925–14945.

Ben-David O, Cohen G, and Fineberg J (2010a) The dynamics of the onset of frictional slip. *Science* 330: 211–214.

Ben-David O, Rubinstein SM, and Fineberg J (2010b) Slip-stick and the evolution of frictional strength. *Nature* 463: 76–79.

Ben-David O and Fineberg J (2011) Static friction coefficient is not a material constant. *Physical Review Letters* 106: 254301.

Beroza G and Ide S (2011) Slow earthquakes and non-volcanic tremor. *Annual Review of Earth and Planetary Sciences* 39: 271–296.

Berthoud P and Baumberger T (1998) Role of asperity creep in time- and velocity-dependent friction of a polymer glass. *Europhysics Letters* 41: 617–622.

Berthoud P, Baumberger T, G'Sell C, and Hiver J-M (1999) State- and rate-dependent friction at a multicontact interface between polymer glasses: I. Static friction. *Physical Review B* 59: 14313–14327.

Bishop AW (1954) Discussion. *Geotechnique* 4: 43–45.

Blanpied ML and Tullis TE (1986) The stability and behavior of a frictional system with a two state variable constitutive law. *Pure and Applied Geophysics* 124: 415–430.

Boettcher MS and Marone C (2004) The effect of normal force vibrations on the strength and stability of steadily creeping faults. *Journal of Geophysical Research* 109: B03406. http://dx.doi.org/10.1029/2003JB002824.

Bocquet L, Charlaix E, Ciliberto S, and Crassous J (1998) Moisture-induced ageing in granular media and the kinetics of capillary condensation. *Nature* 396: 735–737.

Bos B and Spiers CJ (2000) Effect of phyllosilicates on fluid-assisted healing of gouge-bearing faults. *Earth and Planetary Science Letters* 184: 199–210. http://dx.doi.org/10.1016/S0012-821X(00)00304-6.

Bos B and Spiers CJ (2002) Fluid-assisted processes in gouge-bearing faults: Insights from experiments on a rock analogue system. *Pure and Applied Geophysics* 159: 2537–2566. http://dx.doi.org/10.1007/s00024-002-8747-2.

Boulton C, Carpenter BM, Toy V, and Marone C (2012) Physical properties of surface outcrop cataclastic fault rocks, Alpine Fault, New Zealand. *Geochemistry, Geophysics, Geosystems* 13: Q01018. http://dx.doi.org/10.1029/2011GC003872.

Brockley CA and Davis HR (1968) The time-dependence of static friction. *Journal of Tribology* 90: 35–41. http://dx.doi.org/10.1115/1.3601558.

Brown KM, Kopf A, Underwood MB, and Weinberger JL (2003) Compositional and fluid pressure controls on the state of stress on the Nankai subduction thrust. *Earth and Planetary Science Letters* 21: 589–603.

Bhushan B, Israelachvili JN, and Landman U (1995) Nanotribology: Friction, wear and lubrication at the atomic scale. *Nature* 374: 607–616.

Byrne DE, Davis DM, and Sykes LR (1988) Loci and maximum size of thrust earthquakes and the mechanics of the shallow region of subduction zones. *Tectonics* 7: 833–857.

Carpenter BM, Marone C, and Saffer DM (2011) Weakness of the San Andreas Fault revealed by samples from the active fault zone. *Nature Geoscience* 4: 251–254. http://dx.doi.org/10.1038/ngeo1089.

Carpenter BM, Saffer DM, and Marone C (2012) Frictional properties and sliding stability of the San Andreas Fault from deep drill core. *Geology* 40. http://dx.doi.org/10.1130/G33007.1.

Chester FM (1994) Effects of temperature on friction: Constitutive equations and experiments with quartz gouge. *Journal of Geophysical Research* 99: 7247–7261.

Chester FM and Chester JS (1998) Ultracataclasite structure and friction processes of the Punchbowl fault, San Andreas system, California. *Tectonophysics* 295: 199–221.

Chester FM and Higgs NG (1992) Multimechanism friction constitutive model for ultrafine quartz gouge at hypocentral conditions. *Journal of Geophysical Research* 97: 1859–1870.

Collettini C, Niemeijer A, Viti C, and Marone C (2009) Fault zone fabric and fault weakness. *Nature* 462: 907–910. http://dx.doi.org/10.1038/nature08585.

Collettini C, Niemeijer A, Viti C, Smith SAF, and Marone C (2011) Fault zone fabric and fault weakness. *Earth and Planetary Science Letters* 311: 316–327. http://dx.doi.org/10.1016/j.epsl.2011.09.020.

Coulomb CA (1785) Théorie des machines simples, en ayant égard au frotement de leurs parties, et a la roideur des cordages. *Mémoires de Mathematique & de Physique*, X, Paris, pp. 161–342.

Davis D, Suppe J, and Dahlen FA (1983) Mechanics of fold-and-thrust belts and accretionary wedges. *Journal of Geophysical Research* 88: 1153–1172.

den Hartog SAM, Niemeijer AR, and Spiers CJ (2012a) New constraints on megathrust slip stability under subduction zone P–T conditions. *Earth and Planetary Science Letters* 353: 240–252. http://dx.doi.org/10.1016/j.epsl.2012.08.022.

den Hartog SAM, Niemeijer AR, and Spiers CJ (2013) Friction on subduction megathrust faults: beyond the illite-muscovite transition. *Earth and Planetary Science Letters* 373: 8–19.

den Hartog SAM, Peach CJ, De Winter DAM, Spiers CJ, and Shimamoto T (2012b) Frictional properties of megathrust fault gouges at low sliding velocities: New data on effects of normal stress and temperature. *Journal of Structural Geology* 38: 156–171.

den Hartog SAM and Spiers CJ (2013) Influence of subduction zone conditions and gouge composition on frictional slip stability of megathrust faults. *Tectonophysics* 600: 75–90. http://dx.doi.org/10.1016/j.tecto.2012.11.006.

Deng X and Underwood MB (2001) Abundance of smectite and the location of a plate-boundary fault, Barbados accretionary prism. *Geological Society of America Bulletin* 113: 495–507.

Di Toro G, Han R, Hirose T, et al. (2011) Fault lubrication during earthquakes. *Nature* 471: 494–498. http://dx.doi.org/10.1038/nature09.

Di Toro G, Hirose T, Nielsen S, Pennacchioni G, and Shimamoto T (2006) Natural and experimental evidence of melt lubrication of faults during earthquakes. *Science* 311: 647–649.

Dieterich JH (1972) Time dependent friction in rocks. *Journal of Geophysical Research* 77: 3690–3697.

Dieterich JH (1978) Time-dependent friction and the mechanics of stick-slip. *Pure and Applied Geophysics* 116: 790–805.

Dieterich JH (1979) Modeling of rock friction: Experimental results and constitutive equations. *Journal of Geophysical Research* 84: 2161–2168.

Dieterich JH (1981) Constitutive properties of faults with simulated gouge. In: Carter NL, Friedman M, Logan JM, and Sterns DW (eds.) *Mechanical Behavior of Crustal Rocks, AGU Monograph*, vol. 24, pp. 103–120.

Dieterich JH (1992) Earthquake nucleation on faults with rate- and state-dependent friction. *Tectonophysics* 211: 115–134.

Dieterich JH (1994) A constitutive law for rate of earthquake production and its application to earthquake clustering. *Journal of Geophysical Research* 99: 2601–2618.

Dieterich JH and Kilgore BD (1994) Direct observation of frictional contacts: New insights for state-dependent properties. *Pure and Applied Geophysics* 143: 283–302. http://dx.doi.org/10.1007/BF00874332.

Dieterich JH and Kilgore B (1996a) Implications of fault constitutive properties for earthquake prediction. *Proceedings of the National Academy of Sciences* 93: 3787–3794.

Dieterich JH and Kilgore B (1996b) Imaging surface contacts; power law contact distributions and contact stresses in quartz, calcite, glass, and acrylic plastic. *Tectonophysics* 256: 219–239.

Dowson D (1979) *History of Tribology*. New York: Longman.

Edmond JM and Paterson MS (1972) Volume changes during the deformation of rocks at high pressures. *International Journal of Rock Mechanics and Mining Sciences & Geomechanics Abstracts* 9: 161–182.

Fredrich JT and Evans B (1992) Strength recovery along simulated faults by solution transfer processes. *Proceedings of the 33rd U.S. Symposium on Rock Mechanics*, pp. 121–130.

Frye KM and Marone C (2002) The effect of humidity on granular friction at room temperature. *Journal of Geophysical Research* 107(11): 2309. http://dx.doi.org/10.1029/2001JB000654.

Géminard J-C, Losert W, and Gollub JP (1999) Frictional mechanics of wet granular material. *Physical Review E* 59: 5881–5890.

Gettemy GL and Tobin HJ (2003) Tectonic signatures in centimeter-scale velocity–porosity relationships of Costa Rica convergent margin sediments. *Journal of Geophysical Research* 108: 2494. http://dx.doi.org/10.1029/2001JB00073.

Goldsby DL and Tullis TE (2011) Flash heating leads to low frictional strength of crustal rocks at earthquake slip rates. *Science* 334: 216–218.

Gu J-C, Rice JR, Ruina AL, and Tse ST (1984) Slip motion and stability of a single degree of freedom elastic system with rate and state dependent friction. *Journal of the Mechanics and Physics of Solids* 32: 167–196.

Haines SH, van der Pluijm BA, Ikari MJ, Saffer DM, and Marone C (2009) Clay fabric intensity in natural and artificial fault gouges: Implications for brittle fault zone processes and sedimentary basin clay formation. *Journal of Geophysical Research* 114: B05406. http://dx.doi.org/10.1029/2008JB005866.

He G, Müser MH, and Robbins MO (1999) Adsorbed layers and the origin of static friction. *Science* 284: 1650–1652.

Heslot F, Baumberger T, Perrin B, Caroli B, and Caroli C (1994) Creep, stick-slip, and dry-friction dynamics: experiments and a heuristic model. *Physical Review E* 49: 4973–4988.

Hong T and Marone C (2005) Effects of normal stress perturbations on the frictional properties of simulated faults. *Geochemistry, Geophysics, Geosystems* 6: Q03012. http://dx.doi.org/10.1029/2004GC000821.

Hoskins ER, Jaeger JE, and Rosengren KJ (1968) A medium scale direct friction experiment. *International Journal of Rock Mechanics and Mining Sciences* 5: 143–154.

Hyndman RD, Yamano M, and Oleskevich DA (1997) The seismogenic zone of subduction thrust faults. *The Island Arc* 6: 244–260.

Ida Y (1972) Cohesive force across the tip of a longitudinal−shear crack and Griffith's specific surface energy. *Journal of Geophysical Research* 77: 3796–3805.

Ide S, Beroza GC, Shelly DR, and Uchide T (2007) A scaling law for slow earthquakes. *Nature* 447: 76–79. http://dx.doi.org/10.1038/nature05780.

Ikari MJ, Saffer DM, and Marone C (2007) Effect of hydration state on the frictional properties of montmorillonite-based fault gouge. *Journal of Geophysical Research* 112: B06423. http://dx.doi.org/10.1029/2006JB004748.

Ikari MJ, Saffer DM, and Marone C (2009) Frictional and hydrologic properties of clay-rich fault gouge. *Journal of Geophysical Research* 114: B05409. http://dx.doi.org/10.1029/2008JB006089.

Ikari M, Marone C, and Saffer DM (2011) On the relation between fault strength and frictional stability. *Geology* 39: 83–86. http://dx.doi.org/10.1130/G31416.

Johnson KL (1996) Continuum mechanics modelling of adhesion and friction. *Langmuir* 12: 4510–4513.

Johnson KL (1997) Adhesion and friction between a smooth elastic spherical asperity and a plane surface. *Proceedings of the Royal Society of London A* 453: 163–179.

Johnson T (1981) Time dependent friction of granite: implications for precursory slip on faults. *Journal of Geophysical Research* 86: 6017–6028.

Karner SL, Marone C, and Evans B (1997) Laboratory study of fault healing and lithification in simulated fault gouge under hydrothermal conditions. *Tectonophysics* 277: 41–55.

Karner SL and Marone C (1998) The effect of shear load on frictional healing in simulated fault gouge. *Geophysical Research Letters* 25: 4561–4564.

Karner SL and Marone C (2001) Frictional restrengthening in simulated fault gouge: effect of shear load perturbations. *Journal of Geophysical Research* 106: 19319–19337.

Kanagawa K, Cox SF, and Zhang S (2000) Effects of dissolution-precipitation processes on the strength and mechanical behavior of quartz gouge at high-temperature hydrothermal conditions. *Journal of Geophysical Research* 105(B5): 11115–11126.

Kato N, Yamamoto K, Yamamoto H, and Hirasawa T (1992) Strain-rate effects on frictional strength and the slip nucleation process. *Tectonophysics* 211: 269–282.

Kilgore B, Lozos J, Beeler N, and Oglesby D (2012) Laboratory observations of fault strength in response to changes in normal stress. *Journal of Applied Mechanics* 79: 031007.

King DSH and Marone C (2012) Frictional properties of olivine at high temperature with applications to the strength and dynamics of the oceanic lithosphere. *Journal of Geophysical Research* 117: B12203. http://dx.doi.org/10.1029/2012JB009511.

Knuth MW, Tobin HJ, and Marone C (2013) Ultrasonic velocity of sheared granular material: Implications for the evolution of dynamic elastic moduli with shear strain. *Granular Matter* 15: 499–515. http://dx.doi.org/10.1007/s10035-013-0420-1.

Li YG, Vidale JE, Aki K, Xu F, and Burdette T (1998) Evidence of shallow fault zone strengthening after the 1992 M7.5 Landers, California, earthquake. *Science* 279: 217–219.

Li Y-G and Vidale JE (2001) Healing of the shallow fault zone from 1994–1998 after the 1992 M7.5 Landers, California, earthquake. *Geophysical Research Letters* 28. http://dx.doi.org/10.1029/2001GL012922.

Li Q, Tullis TE, Goldsby D, and Carpick RW (2011) Frictional ageing from interfacial bonding and the origins of rate and state friction. *Nature* 480: 233–236.

Linker MF and Dieterich JH (1992) Effects of variable normal stress on rock friction: Observations and constitutive equations. *Journal of Geophysical Research* 97: 4923–4940.

Liu Y (2013) Numerical simulations on megathrust rupture stabilized under strong dilatancy strengthening in slow slip region. *Geophysical Research Letters* 40: 1311–1316. http://dx.doi.org/10.1002/grl.50298.

Losert W, Géminard J-C, Nansuno S, and Gollub J-P (2000) Mechanisms for slow strengthening in granular materials. *Physical Review E* 61: 4060.

Mair K and Marone C (1999) Friction of simulated fault gouge for a wide range of velocities and normal stresses. *Journal of Geophysical Research* 104: 28899–28914.

Marone C (1991) A note on the stress–dilatancy relation for simulated fault gouge. *Pure and Applied Geophysics* 137: 409–419.

Marone C (1998a) Laboratory-derived friction constitutive laws and their application to seismic faulting. *Annual Review of Earth and Planetary Sciences* 26: 643–696.

Marone C (1998b) The effect of loading rate on static friction and the rate of fault healing during the earthquake cycle. *Nature* 391: 69–72.

Marone C and Scholz C (1988) The depth of seismic faulting and the upper transition from stable to unstable slip regimes. *Geophysical Research Letters* 15: 621–624.

Marone C and Scholz CH (1989) Particle-size distribution and microstructures within simulated fault gouge. *Journal of Structural Geology* 11: 799–814.

Marone C and Kilgore B (1993) Scaling of the critical slip distance for seismic faulting with shear strain in fault zones. *Nature* 362: 618–621.

Marone C and Cox SJD (1994) Scaling of rock friction constitutive parameters: the effects of surface roughness and cumulative offset on friction of gabbro. *Pure and Applied Geophysics* 143: 359–386.

Marone C and Saffer DM (2007) Fault friction and the upper transition from seismic to aseismic faulting. In: Dixon TH and Moore JC (eds.) *The Seismogenic Zone of Subduction Thrust Faults*, p. 692. New York: Columbia University Press, ISBN: 978-0-231-13866-6.

Marone C, Raleigh CB, and Scholz CH (1990) Frictional behavior and constitutive modeling of simulated fault gouge. *Journal of Geophysical Research* 95: 7007–7025.

Marone C, Hobbs BE, and Ord A (1992) Coulomb constitutive laws for friction: Contrasts in frictional behavior for distributed and localized shear. *Pure and Applied Geophysics* 139: 195–214.

Marone C, Vidale JE, and Ellsworth W (1995) Fault healing inferred from time dependent variations in source properties of repeating earthquakes. *Geophysical Research Letters* 22: 3095–3098.

Marone C, Cocco M, Richardson E, and Tinti E (2009) The critical slip distance for seismic and aseismic fault zones of finite width. In: Fukuyama E (ed.) *Fault-zone Properties and Earthquake Rupture Dynamics*. *International Geophysics Series*, vol. 94, pp. 135–162. Burlington, MA: Elsevier.

McLaskey GC, Thomas AM, Glaser SD, and Nadeau RM (2012) Fault healing promotes high-frequency earthquakes in laboratory experiments and on natural faults. *Nature* 491: 101–104. http://dx.doi.org/10.1038/nature11512.

Montgomery DR and Jones DL (1992) How wide is the Calaveras fault zone? Evidence for distributed shear along a major fault in central California. *Geology* 20: 55–58. http://dx.doi.org/10.1130/0091-7613.

Moore JC and Saffer DM (2001) Updip limit of the seismogenic zone beneath the accretionary prism of Southwest Japan: An effect of diagenetic to low-grade metamorphic processes and increasing effective stress. *Geology* 29: 183–186.

Muhuri SK, Dewers TA, Scott TE Jr., and Reches Z (2003) Interseismic fault strengthening and earthquake-slip instability: Friction or cohesion? *Geology* 31: 881–884. http://dx.doi.org/10.1130/G19601.1.

Nakatani M and Mochizuki H (1996) Effects of shear stress applied to surfaces in stationary contact on rock friction. *Geophysical Research Letters* 23: 869–872.

Newman AV, Schwartz SY, Gonzalez V, DeShon HR, Protti JM, and Dorman LM (2002) Along-strike variability in the seismogenic zone below Nicoya Peninsula, Costa Rica. *Geophysical Research Letters* 29(20): 1977.

Niemeijer AR and Spiers CJ (2005) Influence of phyllosilicates on fault strength in the brittle-ductile transition: Insights from rock analogue experiments. *Geological Society of London, Special Publication* 245: 303–327.

Niemeijer AR and Spiers CJ (2006) Velocity dependence of strength and healing behavior in simulated phyllosilicate-bearing fault gouge. *Tectonophysics* 427: 231–253. http://dx.doi.org/10.1016/j.tecto.2006.03.048.

Niemeijer AR and Spiers CJ (2007) A microphysical model for strong velocity weakening in phyllosilicate-bearing fault gouges. *Journal of Geophysical Research* 112: B10405. http://dx.doi.org/10.1029/2007JB005008.

Niemeijer AR, Marone C, and Elsworth D (2008) Healing of simulated fault gouges aided by pressure solution: Results from rock analogue experiments. *Journal of Geophysical Research* 113: B04204. http://dx.doi.org/10.1029/2007JB005376.

Niemeijer A, Marone C, and Elsworth D (2010) Frictional strength and strain weakening in simulated fault gouge: Competition between geometrical weakening and chemical strengthening. *Journal of Geophysical Research* 115: B10207. http://dx.doi.org/10.1029/2009JB000838.

Niemeijer A, Di Toro G, Nielsen S, and Di Felice F (2011) Frictional melting of gabbro under extreme experimental conditions of normal stress, acceleration, and sliding velocity. *Journal of Geophysical Research* 116: B07404. http://dx.doi.org/10.1029/2010JB008181.

Noda H and Shimamoto T (2009) Constitutive properties of clayey fault gouge from the Hanaore fault zone, southwest Japan. *Journal of Geophysical Research* 114: B04409. http://dx.doi.org/10.1029/2008JB005683.

Obara K (2002) Nonvolcanic deep tremor associated with subduction in southwest Japan. *Science* 296: 1679–1681.

Ohnaka M, Akatsu M, Mochizuki H, Odedra A, Tagashira F, and Yamamoto Y (1997) A constitutive law for the shear failure of rock under lithospheric conditions. *Tectonophysics* 277: 1–27.

Olsen MP, Scholz CH, and Léger A (1998) Healing and sealing of a simulated fault gouge under hydrothermal conditions: Implications for fault healing. *Journal of Geophysical Research* 103(B4): 7421–7430.

Palmer AC and Rice JR (1973) The growth of slip surfaces in the progressive failure of over-consolidated clay. *Proceedings of the Royal Society of London, Series A* 332: 527–548.

Paterson MS (1995) A theory for granular flow accommodated by material transfer via an intergranular fluid. *Tectonophysics* 245(3–4): 135–151.

Peng Z and Gomberg J (2010) An integrated perspective of the continuum between earthquakes 187 and slow-slip phenomena. *Nature Geoscience* 3: 599–607.

Peng Z, Vidale JE, Marone C, and Rubin A (2005) Systematic variations in recurrence interval and moment of repeating aftershocks. *Geophysical Research Letters* 32: L15301. http://dx.doi.org/10.1029/2005GL022626.

Perrin G, Rice JR, and Zheng G (1995) Self-healing slip pulse on a frictional surface. *Journal of the Mechanics and Physics of Solids* 43: 1461–1495.

Persson BNJ (1998) *Sliding Friction, Physical Principles and Applications*. Berlin/Heidelberg/NewYork: Springer-Verlag, ISBN: 3-540-63296-4.

Rabinowicz E (1951) The nature of static and kinetic coefficients of friction. *Journal of Applied Physics* 22: 1373–1379.

Rabinowicz E (1956) Stick and slip. *Scientific American* 194: 109–118.

Rabinowicz E (1958) The intrinsic variables affecting the stick-slip process. *Proceedings of the Physical Society of London* 71: 668–675.

Ranjith K and Rice JR (1999) Stability of quasi-static slip in a single degree of freedom elastic system with rate and state dependent friction. *Journal of the Mechanics and Physics of Solids* 47: 1207–1218.

Rathbun AP and Marone C (2010) Effect of strain localization on frictional behavior of sheared granular materials. *Journal of Geophysical Research* 115: B01204. http://dx.doi.org/10.1029/2009JB006466.

Rathbun AP and Marone C (2013) Symmetry and the critical slip distance in rate and state friction laws. *Journal of Geophysical Research* 118. http://dx.doi.org/10.1002/jgrb.50224.

Reinen LA, Weeks JD, and Tullis TE (1994) The frictional behavior of lizardite and antigorite serpentinites: Experiments, constitutive models, and implications for natural faults. *Pure and Applied Geophysics* 143: 317–358.

Renard F, Beaupretre S, Voisin C, et al. (2012) Strength evolution of a reactive frictional interface is controlled by the dynamics of contacts and chemical effects. *Earth and Planetary Science Letters* 341: 20–34. http://dx.doi.org/10.1016/j.epsl.2012.04.048.

Rice JR (1983) Constitutive relations for fault slip and earthquake instabilities. *Pure and Applied Geophysics* 121: 443–475.

Rice JR (1993) Spatio-temporal complexity of slip on a fault. *Journal of Geophysical Research* 98: 9885–9907.

Rice JR and Ruina AL (1983) Stability of steady frictional slipping. *Journal of Applied Mechanics* 105: 343–349.

Rice JR, Lapusta N, and Ranjith K (2001) Rate and state dependent friction and the stability of sliding between elastically deformable solids. *Journal of the Mechanics and Physics of Solids* 49(9): 1865–1898.

Richardson E and Marone C (1999) Effects of normal stress vibrations on frictional healing. *Journal of Geophysical Research* 104: 28859–28878.

Roy M and Marone C (1996) Earthquake nucleation on model faults with rate and state dependent friction: the effects of inertia. *Journal of Geophysical Research* 101: 13919–13932.

Rubin AM (2008) Episodic slow slip events and rate-and-state friction. *Journal of Geophysical Research* 113: B11414. http://dx.doi.org/10.1029/2008JB005642.

Rubin AM and Ampuero JP (2005) Earthquake nucleation on (aging) rate and state faults. *Journal of Geophysical Research* 110(B11312). http://dx.doi.org/10.1029/2005jb003686.

Rubinstein SM, Cohen G, and Fineberg J (2004) Detachment fronts and the onset of dynamic friction. *Nature* 430: 1005–1009.

Ruina A (1983) Slip instability and state variable friction laws. *Journal of Geophysical Research* 88: 10359–10370.

Ruina AL (1980) Friction laws and instabilities: A quasistatic analysis of some dry frictional behavior. PhD Thesis, Brown University.

Saffer DM and Marone C (2003) Comparison of smectite and illite frictional properties: Application to the updip limit of the seismogenic zone along subduction megathrusts. *Earth and Planetary Science Letters* 215: 219–235.

Saffer DM, Frye KM, Marone C, and Mair K (2001) Laboratory results indicating complex and potentially unstable frictional behavior of smectite clay. *Geophysical Research Letters* 28: 2297–2300. http://dx.doi.org/10.1029/2001GL012869.

Schleicher A, van der Pluijm B, Solum J, and Warr L (2006) Origin and significance of clay-coated fractures in mudrock fragments of the SAFOD borehole (Parkfield California). *Geophysical Research Letters* 33. http://dx.doi.org/10.1029/2006GL026505.

Scholz CH (1988a) The critical slip distance for seismic faulting. *Nature* 336: 761–763.

Scholz CH (1988b) The brittle-plastic transition and the depth of seismic faulting. In: Zankl H, Delliere J, and Prashnowsky A (eds.) *Detachment and Shear*. Geologische Rundschau 77: 319–328

Scholz CH (1998) Earthquakes and friction laws. *Nature* 391: 37–42.

Scholz CH (2002) *The Mechanics of Earthquakes and Faulting*, p. 439. New York: Cambridge.

Scholz CH, Molnar P, and Johnson T (1972) Detailed studies of frictional sliding of granite and implications for the earthquake mechanism. *Journal of Geophysical Research* 77: 6392–6406.

Scott D, Marone C, and Sammis C (1994) The apparent friction of granular fault gouge in sheared layers. *Journal of Geophysical Research* 99: 7231–7247.

Segall P and Rice JR (1995) Dilatancy, compaction, and slip instability of a fluid infiltrated fault. *Journal of Geophysical Research* 100: 22155–22173.

Shelly DR, Beroza GC, and Ide S (2007) Non-volcanic tremor and low frequency earthquake swarms. *Nature* 446: 305–307.

Shimamoto T and Logan JM (1981a) Effects of simulated clay gouges on the sliding behavior of Tennessee sandstone. *Tectonophysics* 75: 243–255. http://dx.doi.org/10.1016/0040-1951(81)90276-6.

Shimamoto T and Logan JM (1981b) Effects of simulated fault gouge on the sliding behavior of Tennessee Sandstone: Nonclay gouges. *Journal of Geophysical Research* 86: 2902–2914. http://dx.doi.org/10.1029/JB086iB04p02902.

Sleep NH, Richardson E, and Marone C (2000) Physics of friction and strain rate localization in simulated fault gouge. *Journal of Geophysical Research* 105: 25875–25890.

Tadokoro K and Ando M (2002) Evidence for rapid fault healing derived from temporal changes in S wave splitting. *Geophysical Research Letters* 29. http://dx.doi.org/10.1029/2001GL013644.

Tenthorey E, Cox SF, and Todd HF (2003) Evolution of strength recovery and permeability during fluid-rock reaction in experimental fault zones. *Earth and Planetary Science Letters* 206: 161–172.

Tenthorey E and Cox SF (2006) Cohesive strengthening of fault zones during the interseismic period: An experimental study. *Journal of Geophysical Research* 111: B09202. http://dx.doi.org/10.1029/2005JB004122.

Tesei T, Collettini C, Carpenter BM, Viti C, and Marone C (2012) Frictional strength and healing behavior of phyllosilicate-rich faults. *Journal of Geophysical Research* 117: B09402. http://dx.doi.org/10.1029/2012JB009204.

Tse ST and Rice JR (1986) Crustal earthquake instability in relation to the depth variation of frictional slip properties. *Journal of Geophysical Research* 91: 9452–9472.

Tullis TE (1988) Rock friction constitutive behavior from laboratory experiments and its implications for an earthquake prediction field monitoring program. *Pure and Applied Geophysics* 126: 555–588.

Tullis TE (1996) Rock friction and its implications for earthquake prediction examined via models of Parkfield earthquakes. *Proceedings of the National Academy of Sciences of the United States of America* 93: 3803–3810.

Unsworth M, Malin P, Egbert G, and Booker J (1997) Internal structure of the San Andreas fault at Parkfield, CA. *Geology* 25: 359–362.

Vidale JE, Ellsworth WL, Cole A, and Marone C (1994) Variations in rupture process with recurrence interval in a repeated small earthquake. *Nature* 368: 624–626. http://dx.doi.org/10.1038/368624a0.

Vrolijk P (1990) On the mechanical role of smectite in subduction zones. *Geology* 18: 703–707.

Vrolijk P and van der Pluijm B (1999) Clay gouge. *Journal of Structural Geology* 21: 1039–1048.

Yasuhara H, Elsworth D, and Polak A (2003) A mechanistic model for compaction of granular aggregates moderated by pressure solution. *Journal of Geophysical Research* 108(B11): 2530. http://dx.doi.org/10.1029/2003JB002536.

Yasuhara H, Marone C, and Elsworth D (2005) Fault zone restrengthening and frictional healing: The role of pressure solution. *Journal of Geophysical Research: Solid Earth* 110: B06310.

Zheng B and Elsworth D (2012) Evolution of permeability in heterogeneous granular aggregates during chemical compaction: Granular mechanics models. *Journal of Geophysical Research: Solid Earth* 117: B03206.

Zheng B and Elsworth D (2013) Strength evolution in heterogeneous granular aggregates during chemo-mechanical compaction. *International Journal of Rock Mechanics and Mining Sciences* 60.

Zheng G and Rice JR (1998) Conditions under which velocity-weakening friction allows a self-healing versus a cracklike mode of rupture. *Bulletin of the Seismological Society of America* 88: 1466–1483.

4.06 Mechanisms for Friction of Rock at Earthquake Slip Rates

TE Tullis, Brown University, Providence, RI, USA

© 2015 Elsevier B.V. All rights reserved.

4.06.1	**Introduction**	139
4.06.2	**Dynamic Fault-Weakening Mechanisms**	140
4.06.2.1	Normal Stress Reduction or Loss of Contact from Normal Interface Vibrations	140
4.06.2.2	Dynamic Normal Stress Reduction from Elastic or Permeability Mismatch	141
4.06.2.3	Acoustic Fluidization	141
4.06.2.4	Elastohydrodynamic Lubrication	142
4.06.2.5	Thermal Pressurization of Pore Fluid	143
4.06.2.6	Mineral Breakdown-Induced Pressurization of Pore Fluid	145
4.06.2.7	'Flash' Heating/Melting at Asperity Contacts	145
4.06.2.8	Interfacial Lubrication by Friction Melt	148
4.06.2.9	Interfacial Lubrication by Thixotropic Silica Gel	150
4.06.2.10	Rapid Superplastic Deformation of Fine-Grained Material	152
4.06.3	**Friction Resulting from High-Speed Weakening Mechanisms**	153
4.06.4	**Implications of Low Dynamic Friction for Earthquake Stress Drops and for Orientations and Magnitudes of Tectonic Stress**	153
4.06.5	**Conclusions**	154
Acknowledgments		154
References		154

Nomenclature

C	Heat capacity (J kg^{-1} °C)
c_{th}	Thermal diffusivity for flash weakening (mm^2 s^{-1})
D	Diameter of asperity contact (μm)
f	Friction coefficient (dimensionless)
f_0	Friction coefficient at velocities below V_w for flash weakening (dimensionless)
H	Average thickness of the gap in hydrodynamic lubrication (Figure 1) (μm or mm)
L	Wavelength of the asperity in hydrodynamic lubrication (Figure 1) (μm or mm or m)
L^*	Length parameter in thermal pressurization analysis (mm)
p_0	Pore pressure at the initiation of slip (MPa)
T_f	Present background temperature of the fault for flash weakening (°C)
T_w	Weakening temperature for flash weakening (°C)
V	Slip velocity (m s^{-1})
V_w	Velocity at which weakening occurs for flash weakening (m s^{-1})
Λ	Undrained pressurization factor for thermal pressurization (MPa °C^{-1})
ΔH	Amplitude of the asperity in hydrodynamic lubrication (Figure 1) (μm or mm)
α_{hy}	Hydraulic diffusivity (mm^2 s^{-1})
α_{th}	Thermal diffusivity (mm^2 s^{-1})
β_f	Isothermal compressibilities of the fluid (per MPa)
β_n	Pressure expansivity of the pore space (per MPa)
δ	Slip displacement (mm or m)
η	Fluid viscosity (Pa-s)
λ_f	Isobaric volumetric fluid thermal expansivity (per °C)
λ_n	Isobaric volumetric pore space thermal expansivity (per °C)
θ_w	Weakening time for flash weakening (s)
θ	Lifetime of representative contact in flash-weakening analysis (s)
ρc	Specific heat of fault zone material (MPa °C^{-1})
ρ	Material density (kg m^{-3})
σ_n	Normal stress (MPa)
τ_c	Shear stress on asperity contacts for flash weakening (MPa)
τ	Shear stress (MPa)

4.06.1 Introduction

When the first version of this chapter was published in 2007, it began with "Many data exist on the frictional behavior of rocks during slow interseismic, preseismic, and postseismic slip at velocities <1 mm s^{-1}, but only recently have limited data relevant to coseismic slip at velocities ≥ 1 m s^{-1} become available." In the past 7 years, much has changed in this regard – the number of papers reporting experimental data on high-velocity friction has increased manyfold over the number prior to 2007, a pace of research and publication that continues. Consequently, aspects of this chapter will be out of date by the time it is published. To illustrate the growing number of

papers, review papers by Di Toro et al. (2011) and Niemeijer et al. (2012) summarize results from a large number of experimental observations of weakening at high slip velocity. Di Toro et al. (2011) noted that weakening is correlated with the rate of work input, although they also noted that this does not link observed weakening with any particular weakening mechanism. Note that only two new weakening mechanisms, discussed in **Sections 4.06.2.6** and **4.06.2.10**, have been proposed in the 7 years since the first version of this chapter and neither of these has been rigorously documented in publications to date. This testifies to the difficulty of proving what mechanism is responsible for observed experimental behavior and of designing and conducting experiments that unambiguously prove or disprove a theoretically proposed mechanism.

Several high-slip-velocity mechanisms have been proposed that may dramatically reduce the friction coefficient from typical quasi-static values of ~0.6 to values that can approach zero. Rather than list and cite all the experimental observations of high-speed weakening, this chapter is organized around a list of proposed high-speed weakening mechanisms. It briefly discusses these mechanisms, with an emphasis on those for which relevant experimental data exist. Much remains to be learned about all of these mechanisms and their relevance to coseismic slip resistance. Nevertheless, our present understanding suggests that, due to the likely operation of one or more of these mechanisms, coseismic slip resistance is likely to be dramatically lower than would correspond to a coefficient of friction of 0.6. Note that because many processes can potentially be activated at high slip rates, we cannot assume that it is meaningful to extrapolate the small dependence of friction on slip velocity embodied in rate and state friction to seismic slip rates (e.g., Dieterich, 1972, 1978, 1979; Marone, 1998; Ruina, 1983; Tullis, 1988, 1996; Tullis and Weeks, 1986). In addition to the implications for the dynamics of earthquake rupture, the conclusion that dynamic earthquake slip may occur with very low shear resistance also has interesting implications for the magnitudes and orientations of stress in the upper brittle crust. These implications will be briefly discussed later in the text.

The paucity of experimental data on the frictional properties of rocks relevant to coseismic slip in the Earth is due to the extreme conditions of earthquake slip and the difficulty of conducting experiments at these conditions. A laboratory experiment that reproduces the conditions existing during coseismic slip would have to simultaneously involve high slip rates (1–10 m s^{-1}), large slip displacements (0.1–20 m), high effective normal stress (50–200 MPa), elevated pore-fluid pressures (0.4–1 times the normal stress), and elevated temperature (ambient temperatures of 100–300 °C, but as high as 1500 °C in the slipping zone due to shear heating during an earthquake). Although progress is underway toward increasing experimental capability, no apparatus exists that is capable of simultaneously meeting these requirements. Consequently, existing experiments compromise on one or more of these factors.

4.06.2 Dynamic Fault-Weakening Mechanisms

High-speed weakening mechanisms that have been proposed in the literature and are discussed in this chapter are listed in the succeeding text. The four mechanisms marked with an * in the list later in the text are either new additions since the 2007 version of this chapter or ones for which the discussion is significantly altered due to the existence of new experimental data. More minor additions have been made in some of the sections describing the other mechanisms:

1. Normal stress reduction or loss of contact from normal interface vibrations
2. Dynamic normal stress reduction from elastic or permeability mismatch
3. Acoustic fluidization
4. Elastohydrodynamic lubrication
5. Thermal pressurization of pore fluid*
6. Mineral breakdown-induced pressurization of pore fluid*
7. 'Flash' heating/melting at asperity contacts*
8. Interfacial lubrication by friction melt
9. Interfacial lubrication by thixotropic silica gel
10. Rapid superplastic deformation of fine-grained material*.

4.06.2.1 Normal Stress Reduction or Loss of Contact from Normal Interface Vibrations

Brune et al. (1990, 1993) proposed that reductions in normal stress or perhaps even actual loss of contact could occur on a sliding surface during dynamic slip. If this were the case, the shear resistance during slip would be reduced, potentially to zero. That this might occur is supported by direct observations of opening during dynamic slip between foam rubber blocks (Brune et al., 1993). Whether it can occur during sliding on rock surfaces is less clear. Such opening is only predicted theoretically for smooth surfaces in cases where differences in elastic properties exist across the interface (the bimaterial effect discussed later in the text). The observed behavior of the foam rubber occurs without elastic contrast and may result from interactions of the dangling fibers that characterize a cut surface in foam rubber. Intuitively, one might expect that coseismic sliding of fault surfaces with roughness could result in fault opening or reductions in normal stress due to asperities bouncing off one another. Whether this would result in a net reduction in the time- or space-averaged shear resistance is unclear, since the dynamic contact between the asperities would presumably result in both increases in resistance from impact and decreases from rebound. Brune and others (Anooshehpoor and Brune, 1994a,b, 1999; Brown et al., 1991; Brune, 1991, 1996; Brune and Thatcher, 2002; Brune et al., 1990, 1993; Castro et al., 1991) had suggested that unexpectedly large high-frequency *P*-wave radiation observed from earthquakes may indicate that normal vibrations occur during dynamic slip. Brune (2001) had suggested that pulverized rock, which contains no fault-parallel shear features and is found along surface traces of the San Andreas Fault zone, could be the result of such vibrations, although other explanations for pulverized rock exist (e.g., Dor et al., 2006a,b).

This potential weakening mechanism is presently not well understood and whether it is important for reducing the shear resistance during dynamic earthquake slip is unclear. It could be important.

4.06.2.2 Dynamic Normal Stress Reduction from Elastic or Permeability Mismatch

As originally analyzed by Weertman (1963, 1980), if dynamic slip occurs on an interface between two materials having different elastic properties, changes in normal stress on the interface will occur in the slipping area behind the propagating rupture tip. As shown in **Figure 1**, at the end of the rupture that has the less compliant material slipping toward it, the normal stress is increased, whereas at the opposite end of the rupture that has the more compliant material slipping toward it, there is a reduction in normal stress propagating along as a wrinkle-like pulse behind the rupture tip. This can result in two related effects, a reduction in the dynamic resistance to slip and a related tendency for the rupture to propagate at different speeds in the two different directions. In the extreme case, the rupture might propagate unilaterally. Although these effects have been studied extensively (Adams, 1995, 1998, 2001; Andrews and Ben-Zion, 1997; Andrews and Harris, 2005, 2006; Anooshehpoor and Brune, 1999; Ben-Zion, 2001, 2006; Ben-Zion and Andrews, 1998; Ben-Zion and Huang, 2002; Ben-Zion and Shi, 2005; Cochard and Rice, 2000; Harris and Day, 1997, 2005; Ranjith and Rice, 2001; Shi and Ben-Zion, 2006; Weertman, 1963, 1980, 2002; Xia et al., 2005), there has been some controversy about the importance of these effects, especially concerning whether ruptures propagate unilaterally or bilaterally. Thus, some studies (Ben-Zion, 2001, 2006; Ben-Zion and Shi, 2005; Shi and Ben-Zion, 2006) emphasize the tendency for unilateral rupture to occur in bimaterials. Others (Andrews and Harris, 2005, 2006; Harris and Day, 1997, 2005) state that in realistic situations, bilateral rupture will occur in the presence of material contrast, even though the propagation velocity differs in the two directions. There is agreement that if there is no reduction in friction with slip or slip velocity, due, for example, to the other mechanisms discussed in this chapter, then the reduction in normal stress due to elastic mismatch can have significant effects on rupture directivity. However, it appears that if friction decreases significantly with increasing slip and/or slip velocity, then the bimaterial effects become relatively insignificant. Experiments on analogue materials (Xia et al., 2005) show that the ruptures propagate bilaterally, but with different velocities as predicted by theory (e.g., Harris and Day, 1997). Supershear rupture propagation (i.e., in excess of the shear or Rayleigh wave velocity) is also observed (Xia et al., 2005), a result that is not necessarily related to elastic mismatch, but could be due to a large *P*-wave generated by the explosion used to initiate rupture. Supershear propagation is expected if the stress drop is large enough to produce a high-amplitude *P*-wave that induces rupture ahead of the shear wave front (Andrews, 1976). With more normal nucleation processes, the cause of a sufficiently large stress drop and *P*-wave to produce supershear rupture would be a significant decrease in friction with increasing slip and/or slip rate.

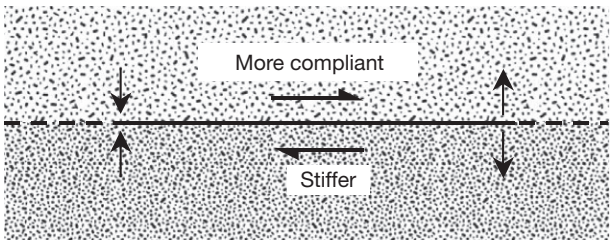


Figure 1 Schematic alterations in normal stress during dynamic rupture across a fault at a bimaterial interface between elastic materials with differing stiffness. Solid line shows section of fault currently slipping or that has slipped in this event; double arrows show fault plane ahead of rupture tip. Near the mode II rupture tips, the normal stress is reduced at the end toward which the more compliant material is slipping and is increased at the end toward which the stiffer material is slipping. Arrows normal to fault show sense of alteration in normal stress due to this bimaterial effect.

Rudnicki and Rice (2006) and Dunham and Rice (2008) discussed a different effect that can arise when a dynamic rupture propagates along a bimaterial interface, that is, the role that differences in permeability on opposite sides of a fault zone can have on the effective normal stress distribution due to changes in pore pressure. Pore pressure is increased in areas of the fault zone where the more permeable material is compressed and reduced where the more permeable material is extended. Since more slip may occur in the areas where the shear resistance is lowered due to increased pore pressure, this mechanism might result in a net reduction in slip-integrated shear resistance. The magnitude of the changes in effective normal stress from this mechanism can be comparable to those due to elastic mismatch (Rudnicki and Rice, 2006). If both the elastic properties and the permeability differ on opposite sides of the fault, the reductions in effective normal stress from one mechanism may either add to or subtract from those due to the other mechanism.

In spite of all of these studies, it is not clear to what extent the frictional resistance during slip is reduced by elastic or permeability mismatch. One might infer that the effects are not large, since elastic mismatch has a smaller effect on the tendency for unilateral rupture propagation in the models than do modest changes in friction with slip (Andrews and Harris, 2005, 2006; Harris and Day, 1997, 2005). Furthermore, the reduction in normal effective stress and hence frictional resistance averaged over the total amount of slip will be less than that that occurs at the rupture tip, since much of the slip may not occur there, where the normal stress is reduced by the maximum amount. In order to determine an average effective reduction in friction, it would be necessary to integrate the product of slip and shear stress as a function of position behind the rupture tip. This does not appear to have been done in dynamic rupture models with bimaterial contrast and further work would be useful in this regard.

In summary, these mechanisms could cause some reduction of dynamic shear resistance in situations where the elastic or permeability mismatch is large enough, but presently, it appears that the effect will be small compared with that potentially available from other weakening mechanisms.

4.06.2.3 Acoustic Fluidization

Melosh (1979) proposed a new mechanism for the reduction of shear resistance inside a deforming granular aggregate, with possible relevance for impact crater collapse, landslides, and earthquakes. Later papers focused more specifically on the application of this mechanism to understanding crater collapse

(Melosh and Ivanov, 1999), landslides (Collins and Melosh, 2003; Melosh and Girdner, 1995), and earthquakes (Melosh, 1996). In this mechanism, termed acoustic fluidization, Melosh envisioned that acoustic waves reverberating inside the shearing aggregate could have sufficient intensity that they could exert large enough forces on the particles that they would be held apart or would touch with less force than would be inferred from the macroscopic stress applied to the boundaries of the aggregate and the real area of contact in the aggregate. If this were the case, then the aggregate could become fluidized and in this case would offer a reduced resistance to shearing. The process is envisioned to be self-sustaining because generation of the acoustic waves continues during shearing and so can balance the losses that will occur due to scattering out of the fault zone and to conversion of elastic energy into heat. Melosh (1996) showed that solutions to the relevant equations exist that allow this mechanism to operate, that they result in an essentially linear viscosity of the fluidized gouge (although later experiments showed highly nonlinear viscosity; Melosh and Girdner, 1995), and that they give shear resistances that are nearly an order of magnitude smaller than the normal stress. However, this theory assumes that pressure fluctuations with frequencies greater than several kilohertz (Melosh, 1979) exist in the acoustic wave field and have magnitudes essentially as large as the normal stress.

While this is an interesting idea, whether such high-amplitude, high-frequency acoustic waves exist is not clear. Measuring them during seismic events would require instrumentation with high-frequency response located within a distance from the fault zone similar to the width of the shearing zone. If current thinking (e.g., see Sibson (2003) and discussion of Rice (2006)) is correct that the active fault shear zone in one seismic event may be thinner than 1–5 mm, perhaps even as thin as 0.2 mm, then following Melosh (1979), the frequencies would have to be in the megahertz range. In this case, the scale at which measurements would have to be made becomes impractical. Indeed, if slip is localized on such a fine scale, it is possible that the granular flow of a fluidized layer of gouge is not even relevant, with slip in the gouge more akin to slip on a plane. On the other hand, if slip is localized in a very narrow gouge zone, then relevant experiments are more feasible than would be the case if the thickness of the shearing zone is on the order of 1 m or more. Results of experiments exploring the relevance of this mechanism have not been published in detail, although some experimental results on masses of sand ~200 mm in dimension seem promising for application to landslides (Melosh and Girdner, 1995; Melosh, personal communication, 2006). More recent relevant very low-pressure experimental work by van der Elst et al. (2012) focuses on the role of both externally imposed and deformation-induced kilohertz frequency vibrations on compaction and its role in determining the thickness of the shearing layer and gives useful references on granular flow.

Although this mechanism cannot be dismissed, at present, its importance for earthquakes is best described as questionable.

4.06.2.4 Elastohydrodynamic Lubrication

Hydrodynamic lubrication is a well-known mechanism for reducing frictional resistance in various machine parts, notably journal bearings (Hamrock, 1994). In these applications, the moving parts are immersed in oil or other fluids. The reduction in frictional resistance between the moving parts occurs because a layer of pressurized fluid separates the parts, preventing solid-to-solid contact. A high fluid pressure can be maintained by a balance between a tendency for pressure decrease due to fluid escape and a tendency for pressure increase due to the geometry of the moving parts. The balance only exists at high rates due to increased hydrodynamic resistance to fluid escape at such rates. Below some speed, the viscous forces that resist the fluid escape are so small that the pressure cannot build up, but at higher speed, the viscous resistance to escape is large and the parts become separated by the pressurized fluid.

Brodsky and Kanamori (2001) proposed a modification to this mechanism that might allow it to operate for a somewhat irregular sliding fault interface. The idea is illustrated in **Figure 2** in which a protrusion is assumed to exist on one side of the fault and a fluid resides in the space between the protrusion and a planar surface. When slip between the surfaces occurs at a sufficiently high velocity, pressure builds up ahead of the protrusion because the gap is closing there and viscous forces impede the fluid escape; this increased pressure tends to reduce the effective normal stress across the interface. In this geometry, the pressure behind the protrusion would be reduced, and so it would initially appear that, averaged over the entire surface, the net effect would be to make no change in the effective normal stress. However, the fluid pressures cause elastic distortion of the solid material and the resulting asymmetry results in a net increase in fluid pressure. Consequently, this process is termed elastohydrodynamic lubrication (Brodsky and Kanamori, 2001).

The geometry illustrated in **Figure 2** is an inadequate view of the interface because it is only two-dimensional. It does show how a noncontacting asperity can lead to an increase in fluid pressure at the interface. However, for this mechanism to operate, many of the asperities must be in actual contact. Otherwise, if a continuous layer of fluid separates the surfaces, the shear resistance would be low without any need for this mechanism to cause weakening. The solution to this is to consider the third dimension, which allows for the presence of an interconnected fluid while at the same time some of the asperities are in solid-to-solid contact. In this case, the resistance could primarily come from these frictional contacts, but the normal and shear forces on them could be reduced by the elastohydrodynamically increased pressure in the interconnected fluid.

What increases in pressure might occur in the interface that could allow elastohydrodynamic lubrication? Brodsky and Kanamori (2001) gave the excess lubrication pressure to be

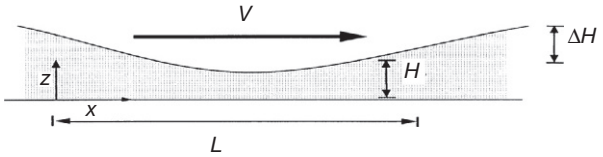


Figure 2 Geometry analyzed for hydrodynamic lubrication, showing moving asperity with closing gap ahead and opening gap behind. Reproduced from Brodsky EE and Kanamori H (2001) Elastohydrodynamic lubrication of faults. *Journal of Geophysical Research* 106: 16357–16374.

$$P_L \approx \frac{6\eta VL\Delta H}{H^3}$$

where H is the average thickness of the gap, ΔH is the amplitude of the asperity, L is the wavelength of the asperity, V is the slip velocity, and η is the viscosity of the fluid (see Figure 2). To illustrate, assuming $V = 1\ \mathrm{m\ s^{-1}}$, $H = 20\ \mu\mathrm{m}$, $L = 1\ \mathrm{mm}$, and $\Delta H = 10\ \mu\mathrm{m}$, we find $P_L = 7.5\ \eta$ MPa when η is given in Pa-s. The viscosity of water at 20 °C is $\sim 10^{-3}$ Pa-s (Sengers and Watson, 1986), so for water, even with gaps as small as 20 μm, $P_L = 0.0075$ MPa, a negligible value compared with values of pore pressure arising from a hydrostatic gradient of ~10 MPa km^{-1}. If the calculation is done for even larger values of L and ΔH, the value of H presumably increases in proportion, and, due to the inverse dependence on H^3, P_L consequently gets even smaller.

What fluids with higher viscosity might allow this mechanism to cause significant lubrication pressures? Brodsky and Kanamori (2001) suggested a slurry of water and fine-grained gouge for which Major and Pierson (1992) measured the viscosity to be in the range of 0.2–7 Pa-s at atmospheric pressure. However, in these experiments, the effective pressure is zero since the fluid pressure is equal to the confining pressure. If this were the case during faulting, then no weakening via this mechanism is needed since the effective normal stress on the fault surface would be zero and consequently the frictional stress would be zero as well. At elevated effective pressures, all rock mechanics experiments in the presence of water suggest that a water-saturated gouge behaves as a Coulomb friction material, not as a viscous fluid (e.g., Byerlee, 1978; Zhang and Tullis, 1998). None of these experiments have been done at extremely high deformation rate, so it is possible that the behavior somehow could change from that of a Coulomb to that of a viscous fluid at high rates. There would not be time for the pressurized pore fluid to equalize throughout the fluid-saturated gouge, but this might only make it act as a Coulomb material at lower effective pressure, not as a fluid. There is certainly no evidence that fluid-saturated gouge does act as a viscous fluid, so the operation of elastohydrodynamic lubrication via a slurry of water and gouge at elevated effective pressure seems unlikely. Presumably, melt with a viscosity of ~10 Pa-s (Spray, 1993) is a viable fluid, as might be the silica gel discussed in a later section. If a layer of relatively weak material such as melt or gel is present, it is unclear whether this mechanism would reduce resistance further. If the weak layer is not very thick, so that it does not completely separate the opposing surfaces, a further reduction in shear resistance by elastohydrodynamic lubrication could be important. On the other hand, parts of the frictional melting experiments of Hirose and Shimamoto (2005) involve only an incomplete melt layer, and during this interval, the frictional resistance is increasing (see discussion in Section 4.06.2.8). The resistance in these experiments seems well described by simple shearing of a viscous layer. In the melt containing experiments of Hirose and Shimamoto (2005) and Tsutsumi and Shimamoto (1997a), over the velocity range investigated, from 0.1 to 1.8 m s^{-1}, no evidence was found of a dramatic decrease with velocity as expected for elastohydrodynamic lubrication (i.e., no Stribeck curve (Brodsky and Kanamori, 2001; Hamrock, 1994) was observed). It appears that few relevant experiments exist, and for the few that do, none demonstrate the operation of this mechanism.

On balance, although it is possible that this mechanism could reduce coseismic frictional resistance, there is presently no reason to believe that it is an important mechanism during earthquakes.

4.06.2.5 Thermal Pressurization of Pore Fluid

In this mechanism, weakening results from transient increases in pore pressure, arising from frictional heat generation and thermal expansion of the pore fluid, resulting in a reduction of the effective normal stress (Andrews, 2002; Goldsby and Tullis, 1997; Lachenbruch, 1980; Lee and Delaney, 1987; Mase and Smith, 1985, 1987; Rempel and Rice, 2006; Rice, 2006; Rice and Cocco, 2007; Sibson, 1973). In order for this mechanism to be effective, the slip must be sufficiently localized that the heat is inputted into a relatively narrow fault zone and rapid enough that there is minimal flow of either heat or pore fluid away from the fault. Observations of exhumed faults (Chester and Chester, 1998; Chester et al., 1993, 2004; Lockner et al., 2000; Noda and Shimamoto, 2005; Sulem et al., 2004; Wibberley and Shimamoto, 2003, 2005) show that at least in some cases, the zone of slip is very narrow, on the order of 100–300 μm (Chester and Chester, 1998), sufficient to meet the localized dissipation requirement. The constraint of low heat flow at seismic slip rates is not difficult to meet, given the low thermal conductivity of rock. However, whether the thermally expanded fluid can pressurize depends critically on the permeability of the fault zone and surrounding country rock. Given the wide range in the permeability of rocks, a wide range of behavior is possible. Existing data on the permeability of rocks in the vicinity of fault zones (Seront et al., 1998; Wibberley, 2002; Wibberley and Shimamoto, 2003) and of experimental faults (Morrow and Lockner, 1997; Morrow et al., 1984, 2001; Zhang and Tullis, 1998; Zhang et al., 1999, 2001) suggest that permeabilities can be low enough to allow significant pressure buildup. Another requirement for this mechanism to be effective is that the dilatancy of the fault is minimal as slip progresses. Otherwise, the pore volume occupied by the fluid could increase due to dilatancy and provide enough space for the thermally expanding fluid that no pressure increase would occur. Some initial dilatancy does occur in rock friction experiments when slip velocity is increased (Beeler and Tullis, 1997; Marone et al., 1990) and implications of this for stability of slip in the absence of frictional heating have been studied (Segall and Rice, 1995). This initial dilatancy has been included in models of thermal pressurization (Rice, 2006), but it is possible that dilatancy may continue with slip due to the nonplanarity of mated fault interfaces (Junger and Tullis, 2003; Power and Tullis, 1992). Further investigation of dilatancy is needed to determine whether it may neutralize thermal pressurization.

The thermal pressurization mechanism is relatively well understood from a theoretical perspective, the most recent analysis being that of Rice (2006). He showed that for slip on a plane and at constant slip velocity V and for constant coefficient of friction f, the shear resistance τ is given by

$$\tau = f(\sigma_n - p_0)\exp\left(\frac{\delta}{L^*}\right)\mathrm{erfc}\left(\sqrt{\frac{\delta}{L^*}}\right)$$

where δ is the slip, σ_n is the normal stress, p_0 is the pore pressure at the initiation of slip at velocity V following any initial dilatancy, and L^* is a length parameter proportional to $1/Vf^2$ with a proportionality constant that contains all of the porothermoelastic properties of the solid and fluid. L^* is given by

$$L^* = \frac{4}{f^2}\left(\frac{\rho c}{\Lambda}\right)^2 \frac{\left(\sqrt{\alpha_{hy}} + \sqrt{\alpha_{th}}\right)^2}{V}$$

where ρc is the specific heat of the fault zone material, α_{hy} and α_{th} are the hydraulic and thermal diffusivities, and Λ is an undrained pressurization factor given by $\Lambda = (\lambda_f - \lambda_n)/(\beta_f - \beta_n)$ where λ_f and λ_n are the isobaric volumetric fluid and pore space thermal expansivities, respectively, β_f is the isothermal compressibility of the fluid, and β_n is the pressure expansivity of the pore space. For a wide range of porothermoelastic parameter values, L^* ranges in magnitude from 2 to 50 mm. The function for τ is such that τ continues to decay with increasing slip, albeit at a decreasing rate, so that the decay cannot be characterized by a slip-weakening length scale. This behavior is shown in **Figure 3**. Rice (2006) presented a 'low-end' and a 'high-end' combination of parameters that result in L^* equal to 4 and 30 mm, respectively, when $V = 1$ m s^{-1} and $f = 0.25$, a value appropriate for flash heating (see discussion later in the text). With these parameter values, the full displacement scales in the four panels of **Figure 3** range from 0.4 mm to 0.4 m for the 'low-end' parameters and from 3 mm to 3 m for the 'high-end' parameters, suggesting that the resistance to slip should decrease to 0.05 times its initial value within 0.4–3 m of slip. Since this is for an initial value of the friction coefficient for flash heating of $f = 0.25$, the effective friction coefficient (τ/σ_n) would be as low as 0.01 after 0.4–3 m of slip. In the absence of weakening by flash heating, the amount of frictional heating would be greater, so the displacement needed to attain a given amount of weakening is smaller. Thus, if the initial friction coefficient $f = 0.6$, the above displacements of 0.4–3 m would be smaller by a factor of 0.17 (Rice, 2006), namely, 68–510 mm for the low-end and high-end parameters, and the effective friction coefficient would have decreased to 0.03.

The 2007 version of this chapter stated, "Experimental confirmation of this mechanism has yet to be accomplished, although it is amenable to such study." This situation is changing. Weakening observed in an initial attempt at exploring this mechanism in the laboratory (Goldsby and Tullis, 1997) was later found to be due to the gel-weakening mechanism described in a later section. In 2007, this chapter said, "However, since some rocks do not show weakening due to gel, it should be possible to revisit the thermal pressurization problem experimentally (Tullis et al., 2007)." This has recently been done.

The most convincing experimental evidence for thermal pressurization has to date only been presented in a USGS Final Technical Report (Tullis and Goldsby, 2013). The experiments were done in the only existing rotary shear apparatus in which a jacketed sample is subject to an elevated confining pressure and a separately controlled pore pressure. In the case of the most interesting experiment, which was done on diabase, the confining pressure was 100 MPa, the normal stress on the sliding surface was 103 MPa, and the initial pore pressure of water was 10 MPa, creating an initial effective normal stress of 93 MPa. Sliding took place on an initially bare surface on which little gouge developed. The permeability of the sample was 6.1×10^{-19} m^2, a low enough value that at the sliding rate of 4.8 mm s^{-1}, any H_2O in which pore pressure increases due to thermal pressurization cannot readily escape. Unfortunately, this also means that it is not possible to directly measure the pore pressure during slip and it needs to be inferred from the change in shear stress, assuming the coefficient of friction is unchanged during sliding. As shown in **Figure 4**, somewhat revised from the figure shown in Tullis and Goldsby (2013), upon a large increase in sliding velocity a decrease in shear strength consistent with thermal pressurization occurs. A similar behavior is shown in another experiment on the San Andreas Fault gouge obtained from the SAFOD borehole (Tullis and Goldsby, 2013). The only viable explanation for the initial decrease in shear stress, especially for the diabase sample, is thermal pressurization.

Other authors have observed reductions in shear stress that might be due to thermal pressurization (De Paola et al., 2011; Faulkner et al., 2011; Kitajima et al., 2010, 2011; Mizoguchi et al., 2009; Ujiie et al., 2013). However, their experiments have all been done in rotary shear experiments in which there is no confining pressure or separately controlled pore pressure and the jackets are simple Teflon sleeves. Thus, it is difficult for any pore pressure to be maintained since the fluid can easily escape the sliding interface or shearing gouge layer and the Teflon sleeve. Transient increases in pore-fluid pressure could possibly develop in this situation, but it is difficult to determine its magnitude or whether it has any influence on the shear resistance. An analysis of the situation by Mizoguchi

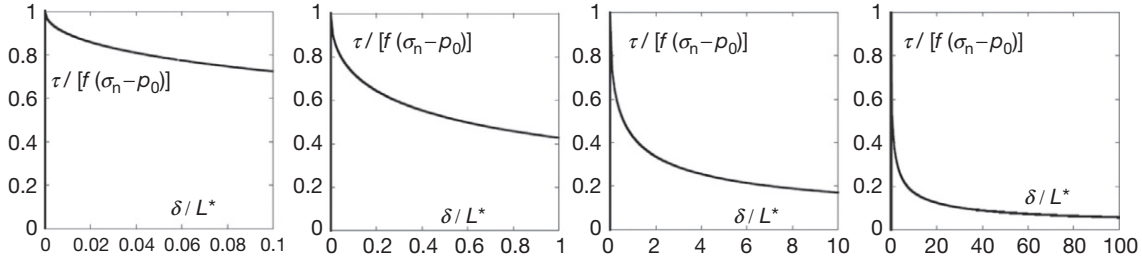


Figure 3 Progressive weakening with slip due to thermal pressurization resulting from slip on a plane at constant slip velocity with constant coefficient of friction. Successive frames show that a characteristic decay slip distance does not exist for this mechanism. See text and Rice, 2006 for meaning and values of L^*. Reproduced from Rice JR (2006) Heating and weakening of faults during earthquake slip. *Journal of Geophysical Research: Solid Earth* 111(B5). http://dx.doi.org/10.1029/2005JB004006.

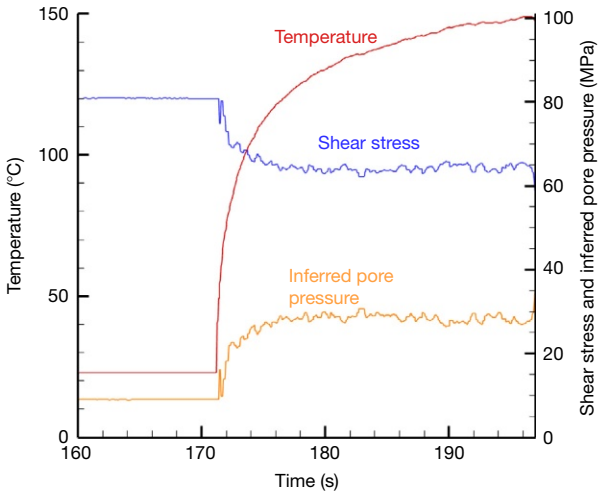


Figure 4 Plots of shear stress, inferred pore-fluid pressure, and temperature for an experiment on diabase bare surfaces showing weakening behavior. The weakening begins immediately upon increasing the sliding velocity from 10 μm s^{-1} to 4.8 mm s^{-1} and we infer it is due to thermal pressurization. Due to low permeability of the sample, the pore-fluid pressure cannot be measured directly and is inferred assuming that all changes of shear stress are due to changes in fluid pressure, that is, assuming the effective stress law holds for the experiment. The temperature is calculated from a finite element model that includes the geometry and material properties of the sample and sample assembly and the time-dependent measured values of shear stress. Zero time on these plots is arbitrary and corresponds to when the temperature calculations began. The leveling out of the shear stress about 6 s after the velocity step occurs because the steel sample grips prevent the temperature from rising as much as it would in an infinite half-space.

et al. (2009) tentatively suggests that the weakening in such experiments is not due to thermal pressurization because the characteristic time for escape of the pore fluid is much shorter than the duration of the weakening, although more data are needed on the values of the permeability and porosity of the gouge during large rapid shearing. A careful analysis of the possibility that thermal pressurization might explain the observed weakening in their experiments was conducted by Kitajima et al. (2011), but uncertainties in the values of the permeability of the gouge layer and the Teflon jacket and how these might vary during the experiment make it difficult to come to a definitive conclusion concerning the role of elevated pore pressure. The most convincing evidences for thermal pressurization in experiments of this type are those of Faulkner et al. (2011) and of Ujiie et al. (2013). Both of these studies showed that unconfined clay-rich samples sheared at about 1 m s^{-1} while surrounded by simple Teflon sleeves had a lower frictional resistance when using low-permeability rock-forcing blocks adjacent to the clay than when using higher-permeability forcing blocks. This suggests that higher pore pressure developed in the samples with low-permeability surroundings and that thermal pressurization from the elevated temperatures computed to result from the rapid sliding is a likely cause.

This mechanism remains a viable one for significant reductions in strength of faults during earthquakes. More study is needed to investigate whether the theoretical predictions are borne out by more experiments and to what extent dilatancy may counteract the predicted increase in fluid pressure. If dilatancy is unimportant, with sufficient slip, the shear resistance due to thermal pressurization can approach zero as shown in **Figure 3** from Rice (2006) and, for example, by Mase and Smith (1985).

4.06.2.6 Mineral Breakdown-Induced Pressurization of Pore Fluid

Another way that elevated pore-fluid pressures could be produced by high-speed slip is that pore fluids such as H_2O or CO_2 could be produced by the thermal breakdown of minerals having appropriate chemical constituents (Brantut et al., 2008, 2010), although it is not clear whether this mechanism caused the weakening observed in the experiments of Brantut et al. (2008). It might initially seem that this would provide a limited source of pressurized fluid since breakdown would deplete the mineral phases breaking down. However, theoretical modeling by Platt and Rice (2013) shows that it is possible for the shearing zone to migrate into adjacent unreacted material during progressive slip and thereby greatly increase the amount of material available for production of pressurized fluids.

It is somewhat unclear whether experimental evidence for the operation of this mechanism currently exists, although it still seems a viable mechanism. Even if it is important in nature, it is difficult to demonstrate in unconfined experiments. The problem to date is that in the high-speed rotary shear experiments done to test this mechanism, the fluid escapes from the unconfined samples, just as in the thermal pressurization case. Experimental observations possibly supporting this mechanism for the breakdown of carbonates to produce pressurized CO_2 were presented by De Paola et al. (2011), Han et al. (2007), and Brantut et al. (2011) for gypsum. However, Han and coworkers subsequently did experiments on the reaction products in the absence of pressurized fluids and found them to be as weak at high slip rates as the carbonates after they decomposed, leaving no weakening to be explained by elevated fluid pressures (Han et al., 2010, 2011). They then attributed the weakening to 'nanoparticle lubrication,' a putative mechanism discussed later in this chapter. A similar situation appears to exist for the experiments done by others that involve liberation of a fluid phase by shear heating. However, some interesting arguments are presented by De Paola et al. (2011) that suggest that it is likely that increases in fluid pressure of CO_2 caused by thermal decomposition did in fact contribute to weakening in their experiments.

Although it is not easy to demonstrate the operation of this mechanism in unconfined experiments and more work needs to be done to understand it better, it could be important for reducing the shear resistance during dynamic earthquake slip.

4.06.2.7 'Flash' Heating/Melting at Asperity Contacts

Contact across frictional interfaces occurs over only a fraction of the nominal area of contact, and conditions at the contact spots are quite different from those on the surface as a whole.

The generally accepted explanation for friction, namely, the linear increase of shear resistance to sliding with increases in normal force, is that the total real area of contact increases in proportion to increases in the nominal normal stress. The real area of contact occurs where high points on the opposing surfaces called asperities touch each other, and the shear resistance to sliding arises from the average shear strength of the contacts τ_c, times the total real area of contact. The average normal stress at the contacts is much higher than the nominal normal stress by the ratio of the nominal contact area to the real area of contact. It is often assumed that the normal stresses at the contacting asperities are equal to the compressional yield strength of the material, although in hard materials, the contacts may be elastic and not reach the yield stress (Johnson, 1985). During sliding, the contacts are short-lived as high points on each surface move past one another, but the number and size of contacts in the population are independent of slip at constant slip velocity, assuming no fracture/comminution occurs. The lifetime θ of a representative contact is equal to the length D of the contact in the slip direction, divided by the slip velocity V.

Locally elevated temperatures occur at contacts and can result in degradation of contact strength and therefore reduction in macroscopic friction. The rate of heat dissipation at a point within the contact is the product of V and the contact shear stress τ_c. Because τ_c is much larger than the nominal macroscopic shear stress (by the ratio of the nominal area of the surface to the real area of contact), the product $\tau_c V$ is much larger than the average heat dissipation rate on the frictional interface. If V is sufficiently high, the local temperature at the contacts is much greater than the average temperature on the surface. Such elevated temperatures were measured for dissimilar metals in sliding contact using the thermocouple effect (Bowden and Ridler, 1936) and for transparent materials by studying visible (Bowden et al., 1947) and infrared (Bowden and Thomas, 1954) radiation. Macroscopic weakening occurs due to thermal degradation of contact strength, and at sufficiently high V, melting may occur. The process is often called 'flash' heating, after small 'stars of light' observed during sliding of transparent materials (Bowden et al., 1947). In cases where melting occurs, the process is called 'flash' melting.

Theoretical analyses of flash heating and weakening have been undertaken by many authors (Archard, 1958; Beeler and Tullis, 2003; Beeler et al., 2008; Blok, 1937; Bowden and Persson, 1960; Bowden and Ridler, 1936; Bowden and Thomas, 1954; Ettles, 1986; Jaegar, 1942; Lim and Ashby, 1987; Lim et al., 1989; Molinari et al., 1999; Rice, 1999, 2006). Rice (1999) conducted an analysis of the weakening velocity and how the steady-state frictional resistance should depend on slip velocity at velocities above the weakening velocity. In Rice's analysis (Rice, 1999), since expanded (Rice, 2006), once a part of an asperity exceeds a weakening temperature T_w, it has zero strength. He obtained the weakening time θ_w, that is, the time that the oldest part of an asperity contact exists before it reaches the weakening temperature, by equating the heat input (i.e., the heat input rate $\tau_c V$ times the age of the oldest part of the asperity $\theta = D/V$) to the thermal energy storage $\rho c(T_w - T_f)$ times an effective distance $(\pi c_{th}\theta)^{1/2}$. That effective distance is obtained from a one-dimensional heat flow analysis, where ρ is the material density, c is its heat capacity, c_{th} is the thermal diffusivity, and T_f is the present average temperature of the fault. Solving this for V and regarding it as the velocity V_w at which weakening first occurs, Rice obtained

$$V_w = (\pi c_{th}/D)[\rho c(T_w - T_f)/\tau_c]^2$$

Beeler (Beeler and Tullis, 2003; Beeler et al., 2008) extended Rice's (Rice, 1999) analysis to include a distribution of contact sizes, allowed the weakened part of the contact to have nonzero strength, and showed how the strength evolves with slip over small slip distances on the order of the contact spot size. Beeler et al. (2008) showed that Rice's $1/V$ dependence of friction on V above the weakening velocity V_w results from variations with velocity of the contributions of the weakened older parts of a typical asperity contact (those parts older than θ_w) and the unweakened younger parts. The result of these analyses is that the friction coefficient $f = f_0(V_w/V)$ where f_0 is the friction coefficient at velocities below V_w. Appropriate values for all of the parameters in Rice's expression for V_w are not well known. The better-known values are c_{th} (1–2 $mm^2\ s^{-1}$) and the specific heat per unit volume ρc (3–4 $MJ\ m^{-3}\ °C^{-1}$). Less well known are D, which depends on the roughness of the contacting surfaces; T_w, which depends on which process causes weakening (thermal softening or melting); and τ_c, which depends on whether the contacts are at their compressive yield strength and whether their shear strength depends on flow or fracture of the crystalline material, on melting of the contact, or on the strength of some other weak layer coating the asperities. Reasonable parameter values are $D = 5$–$20\ \mu m$, $(T_w - T_f) = 1000$–$1700°\ K$, and $\tau_c = 5$–10 GPa. While acknowledging the uncertainty involved, a frequently employed estimate for V_w is $0.1\ m\ s^{-1}$ (Rice, 2006).

As discussed later in the text, data from two independent experimental studies are consistent with flash weakening, although further work is required to be more confident of the underlying mechanism. Verification of this mechanism can be accomplished by conducting experiments at elevated temperature to see if shifts in the predicted V_w with changes in $(T_w - T_f)$ match observed shifts in V_w with changes in $(T_w - T_f)$. Similarly, varying D by varying surface roughness should yield predictable shifts in V_w, according to Rice's expression for V_w in the preceding text.

In unconfined rotary shear experiments by Goldsby (Goldsby and Tullis, 2003, 2011; Tullis and Goldsby, 2003), an abrupt weakening is seen at slip velocities above $0.1\ m\ s^{-1}$, the independently predicted theoretical value of V_w. The weakening occurs after a small slip displacement of about 1 mm. That this is a larger value than contact spot sizes is expected due to the inertia of the testing machine. After the velocity is reduced below $0.1\ m\ s^{-1}$, frictional strength returns to its typical quasi-static value within about 0.01 s (~2 mm of slip). These results are shown in **Figures 5** and **6** for quartzite. Similar behavior has been seen for granite, gabbro, a pure feldspar rock, a fine-grained pure quartz rock, novaculite (Goldsby and Tullis, 2011; Roig Silva et al., 2004b), serpentine (Hirose and Bystricky, 2007; Kohli et al., 2011), and calcite marble (Han et al., 2007). Most attempts to observe textural evidence on the slip surfaces that might support the flash weakening mechanism (e.g., melted asperity tips) have revealed nothing diagnostic, although such evidence might be difficult to locate. However,

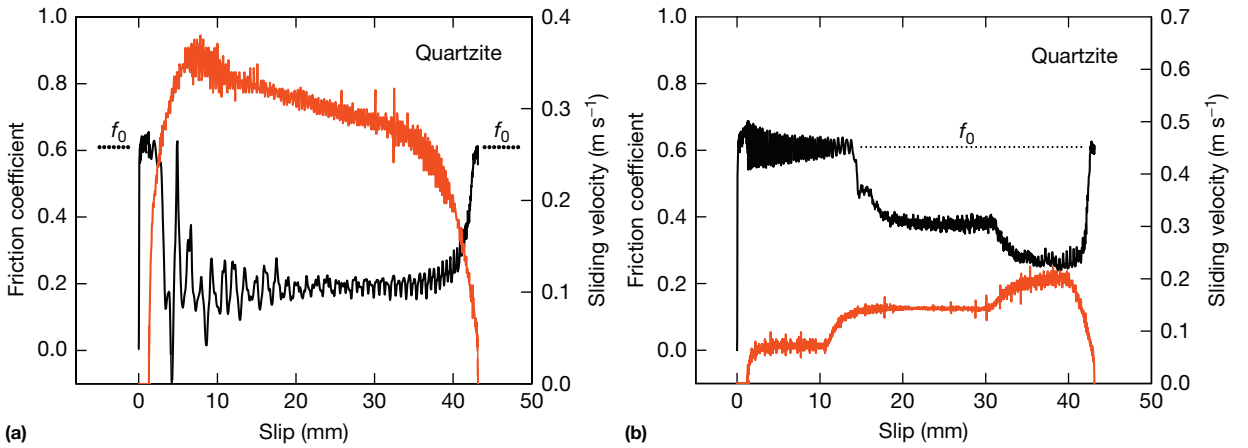


Figure 5 (a) Friction coefficient (black trace) and sliding velocity (red trace) plotted against sliding displacement, for a high-speed continuously varying velocity (CCV) friction experiment on quartzite. The friction coefficient at low slip rates, f_0, has nearly identical values before and after sliding at rapid slip rates. (b) Friction coefficient (black trace) and sliding velocity (red trace) plotted against slip for a velocity stepping test (V_s) test on quartzite. As in (a), the friction coefficient at low slip rates, f_0, has nearly identical values before and after rapid slip. In both tests, upon acceleration, weakening is not observed until a slip of ~3 mm has accrued above $V = 0.10$ m s^{-1}. Reproduced from Figure 1 of Goldsby DL and Tullis TE (2011) Flash heating leads to low strength of crustal rocks at earthquake slip rates. *Science* 334: 216–218. http://dx.doi.org/10.1126/science.1207902.

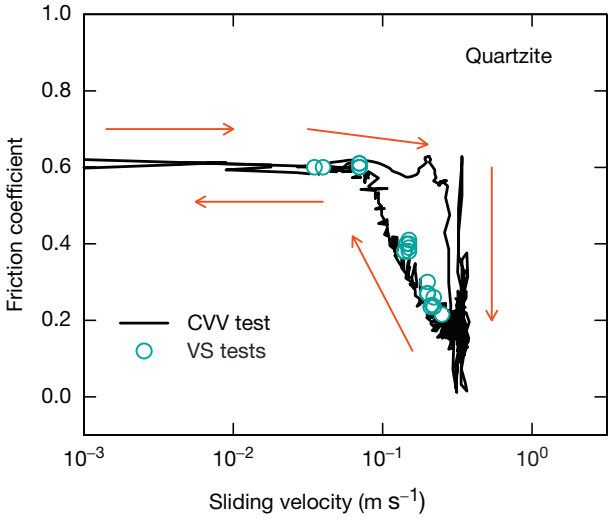


Figure 6 Friction coefficient versus sliding velocity from the CVV test (solid black line) of **Figure 5** and several subsequent VS tests (open blue-gray circles) on the same quartzite sample. Arrows indicate the sequence in which the data were acquired in the CVV test. A hysteresis in the friction coefficient from the CVV experiment results from higher friction on acceleration compared with lower friction on deceleration, attributed to the effects of shear localization and delocalization above and below, respectively, the weakening velocity V_w of about 0.1 m s^{-1}. As shown by Goldsby and Tullis (2011), the friction upon deceleration and the VS test friction follow the predictions of flash weakening theory (Beeler and Tullis, 2003; Beeler et al., 2008; Rice, 2006). Reproduced from Figure 2 of Goldsby DL and Tullis TE (2011) Flash heating leads to low strength of crustal rocks at earthquake slip rates. *Science* 334: 216–218. http://dx.doi.org/10.1126/science.1207902.

in the case of serpentine, small areas of talc have been seen on the surface that seem to verify that a weakening reaction occurred at asperities that experienced flash heating (Kohli et al., 2011). All the characteristics of the mechanical behavior agree with expectations for flash weakening, as the fits to the data shown in **Figure 6**. Further tests in which D and T_f are varied would ideally allow a simple test of whether flash weakening is responsible for the observed behavior. However, experiments done that vary T_f showed that V_w increased with increases in T_f, contrary to the apparent predictions of the previously mentioned equation for V_w. The explanation is that the values of some parameters (thermal conductivity, heat capacity, and contact stress) in the equation also depend on T_f in such a way that V_w is increased (Goldsby et al., 2011; Passelegue et al., 2014).

Two other experimental techniques new to the geophysical research community yield low friction for rocks consistent with flash heating at small displacements and slip rates above the theoretically predicted V_w (Prakash and Yuan, 2004). Results for novaculite are shown in **Figures 7** and **8**. In **Figure 7**, data from a pressure–shear experiment are shown, whereas in **Figure 8**, data from a torsional Kolsky bar experiment are shown. A description of the pressure–shear and torsional Kolsky bar techniques is given by Prakash and Clifton (1993) and Rajagopalan and Prakash (1999), respectively. In both types of experiments, dramatic reductions in shear resistance occur with small slip displacements, to levels consistent with flash weakening. In these pilot experiments, designed in part to demonstrate the feasibility of using these dynamic testing techniques on fully dense rocks, there was no systematic attempt to vary the sliding velocity or to determine the dependence of friction on velocity. However, the techniques show promise for investigating weakening at high slip speeds (1–50 m s^{-1}) and modest slip displacements (0.5–5 mm).

Although flash weakening needs to be studied in more detail in the laboratory, it appears to be a strong candidate for causing abrupt and dramatic weakening during earthquakes. Under conditions in the upper crust where frictional behavior predominates and consequently resistance to earthquake slip occurs over small real areas of contact on fault surfaces, it seems likely that this mechanism would be an important one.

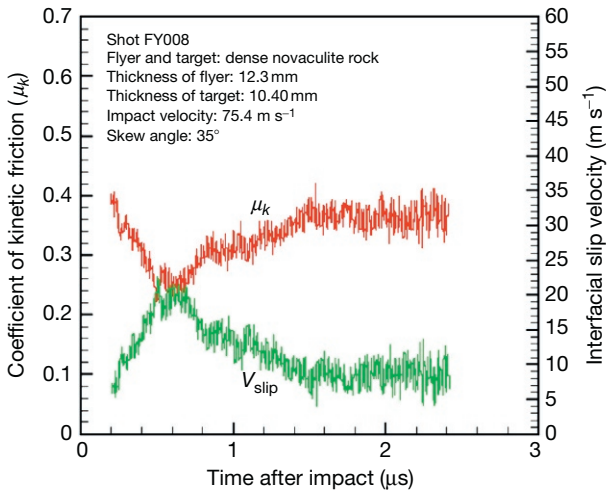


Figure 7 Friction (red) and slip velocity (green) versus time for pressure shear experiment on novaculite. Reproduced from Prakash V and Yuan F (2004) Results of a pilot study to investigate the feasibility of using new experimental techniques to measure sliding resistance at seismic slip rates. *EOS*, *Transactions*, *American Geophysical Union* 85(47). Fall Meeting Supplement, Abstract T21D-02.

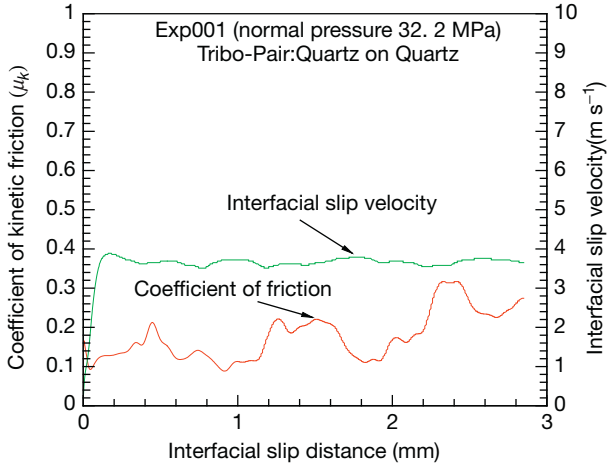

Exp001 (normal pressure 32. 2 MPa)
Tribo-Pair:Quartz on Quartz
Interfacial slip velocity
Coefficient of friction
Coefficient of kinetic friction (μk)
Interfacial slip velocity(m s−1)
Interfacial slip distance (mm)

Figure 8 Friction and slip velocity for torsional Kolsky bar experiment on novaculite. Reproduced from Prakash V and Yuan F (2004) Results of a pilot study to investigate the feasibility of using new experimental techniques to measure sliding resistance at seismic slip rates. *EOS*, *Transactions*, *American Geophysical Union* 85(47). Fall Meeting Supplement, Abstract T21D-02.

Two factors, however, might limit the significance of the flash weakening mechanism. One concerns the magnitude of the shear stress at contacting asperities and the other concerns the thickness of the fault zone across which the shear displacement takes place. If clay minerals dominate the fault zone, the shear stress τ_c might be considerably smaller than 5–10 GPa, increasing V_w to slip rates above those experienced seismically. Also, for flash heating calculations, it is often assumed that τ_c is equal to the theoretical strength, which is on the order of a mineral's shear modulus divided by ~30 (Cottrell, 1965; Hull, 1965), and this seems to be borne out by some experiments (Dieterich and Kilgore, 1994, 1996). On the other hand, in strong brittle materials, frictional contacts can elastically deform rather than plastically yield (Johnson, 1985), so that τ_c could be smaller than often assumed. The other factor is that, even with τ_c approaching the theoretical strength for most silicates, if shearing is distributed over many contacts in a sufficiently thick deforming layer of fault gouge, the slip velocity at every contact in the gouge could be less than V_w even if the total slip velocity across the zone was 1 m s^{-1} or higher. For example, if shearing occurred in a deforming gouge layer only ten grains thick, the slip velocity at each contact might be reduced to 0.1 m s^{-1} even with a total slip velocity of 1 m s^{-1}, which might defeat the mechanism (Beeler et al., 2008).

Both the role of τ_c and the role of distributed shearing can be studied experimentally, one by rapid shearing clay minerals and the other by using gouge layers of varying thickness. Although it is notable that slip often tends to become very localized in gouge experiments (Beeler et al., 1996; Scruggs and Tullis, 1998), as can also occur on natural faults (Chester and Chester, 1998; Chester et al., 1993, 2004; Lockner et al., 2000; Noda and Shimamoto, 2005; Sulem et al., 2004; Wibberley and Shimamoto, 2003), localization is typically less narrow than ten grain diameters (see discussions of factors affecting localization by Rice (2006) and by Rempel and Rice (2006)). How localized the deformation is in the experiments of Goldsby and Tullis (2011) is presently unclear. The experiments involve initially bare rock surfaces and the assumption is that gouge is not generated in significant thicknesses if the initial roughness is low. However, some powdered rock escapes from the edges of the ring-shaped sliding surface. How much generated gouge separates the bare surfaces in these experiments during sliding and how this might affect the observed weakening are unknown and warrant further study. The idea that the presence of gouge may increase the flash-weakening velocity, perhaps beyond velocities attained in experiments or during earthquakes, is supported by experiments of Goldsby (Goldsby and Tullis, 2007) who found that if rougher than normal bare surfaces were used, a thicker layer of gouge was formed and, although flash weakening was found in the same material with less gouge present, no weakening was seen with thicker gouge.

4.06.2.8 Interfacial Lubrication by Friction Melt

If sufficiently large slip occurs at high slip velocity and with high frictional resistance, the temperature of the entire fault surface or zone can become high enough that a layer of melt develops and lubricates the entire fault surface. Melting caused by faulting was suggested as a dynamic fault-weakening mechanism by Jeffreys (1942) and was suggested as an explanation for pseudotachylites by Goldschmidt (1943). The origin and significance of pseudotachylites and/or the reduction of shear resistance by shear melting has been discussed by many other authors (e.g., Bjornerud and Magloughlin, 2004; Di Toro and Pennacchioni, 2004, 2005; Di Toro et al., 2005a,b, 2006; Ferre et al., 2005; Fialko, 2004; Fialko and Khazan, 2005; Kanamori and Heaton, 2000; Nielsen, et al., 2008, 2010; Maddock, 1986; Magloughlin, 1989, 1992; Magloughlin and Spray, 1992; Mase and Smith, 1987; McKenzie and Brune, 1972; Philpotts, 1964;

Rempel and Rice, 2006; Sibson, 1975; Spray, 1989a,b, 1992, 1995, 2005; Spray and Thompson, 1995; Swanson, 1988, 1990, 1992; Wenk, 1978, 1979; Wenk et al., 2000). Two special volumes of *Tectonophysics* were devoted to these rocks (1992, vol. 204, and 2005, vol. 402). The most recent and thorough theoretical treatments are those of Fialko and Khazan (2005) and Nielsen et al. (2008, 2010). In the absence of fluids or for a fluid-saturated fault of high permeability, theory predicts a progression with slip, beginning with flash heating at asperity contacts, to a situation where the overall temperature of the fault T_f gradually increases to the point that it reaches the melting temperature. With continued input of heat, the heat of fusion is provided to more and more material in the fault zone, resulting in a layer of melt along the fault surface. Continued slip results in viscous shearing of this melt layer. Although the melt layer may be significantly weaker than the initial frictional interface, thus lowering the heat input rate, continued heat may be generated by viscous dissipation in the shearing layer, perhaps along with frictional heating occurring where high points on opposite sides of the interface project through the melt layer and contact. It is known that such frictional melting occurs during some earthquakes because pseudotachylites are observed on some faults, although they are often believed to be less abundant than expected (Sibson and Toy, 2006a,b). Questions have been raised concerning whether all pseudotachylites are the product of melting (Wenk, 1978), so it is best to reserve the term for rocks containing definite evidence of melting (such as the presence of microlites, spherulites, clasts with embayments, and amygdules) and use another term such as ultracataclasites for rocks with similar macroscopic morphology and occurrence but for which definite evidence of melting cannot be demonstrated (even though some of them may also have resulted from melting).

Although much more experimental work needs to be conducted, a series of informative experiments on shear melting have been conducted by Spray (1987, 1988, 1993) and by Shimamoto and colleagues (Di Toro et al., 2006; Hirose, 2002; Hirose and Shimamoto, 2003, 2005; Hirose et al., 2000; Lin and Shimamoto, 1998; Shimamoto and Tsutsumi, 1994; Shimamoto and Lin, 1994; Tsutsumi and Shimamoto, 1994, 1996, 1997a,b). An analysis of the observed eventual weakening in the experiments of Hirose and Shimamoto (2005) has been provided by Fialko and Khazan (2005). It is apparent from these experiments that shear melting indeed occurs at sufficiently high slip velocity and shear stress. Because these experiments have all been done without confining pressure, the normal stress is small to prevent fracturing. Therefore, the frictional shear stress prior to melting is lower than that for earthquake-generating faults. Thus, the slip displacements in the experiments are higher (up to 174 m) than those during earthquakes in order to generate sufficient heat and raise the temperature high enough to cause melting. Since the samples are unconfined, it is not possible for any melt pressure to build up in the samples. In fact, the melt is thrown radially out of the rotary-shear-configuration samples by centrifugal force. This resembles the situation often seen in occurrences of pseudotachylites in which only thin layers of pseudotachylite are found along the pseudotachylite generation zones along the faults and reservoirs of injected pseudotachylite are found in veins connected to the generation zones.

Study of the evolution of shear resistance in these experiments (Hirose and Shimamoto, 2005) shows a complex sequence that can be related to the evolution of the microstructures in the samples. The initial mechanical response is a weakening that appears to be due to flash melting, followed by a strengthening that appears to be due to the development and growth of melt patches on the sliding surface, which result in a larger real area of contact than is the case for frictional sliding prior to the development of melt patches. Significant viscous resistance apparently occurs at the melt patches. The final development is a gradual weakening, which accompanies the creation of a continuous melt layer, and continues with the growth in thickness of the melt layer (Hirose and Shimamoto, 2005) and an increase in the melt fraction in the melt–clast mixture on the fault with increasing temperature (Fialko and Khazan, 2005). This sequence of correlated mechanical and textural developments is shown in **Figure 9**.

The shear resistance with a layer of melt is not as low in the experiments in Shimamoto's rotary shear apparatus (Hirose and Shimamoto, 2005; Tsutsumi and Shimamoto, 1997a) as one might expect, that is, the coefficient of friction at what appears to be nearly steady state is 0.6, a value typical of rock friction without melt in experiments with slip velocities of 1 mm s^{-1} or lower. The initial weakening shown in **Figure 9**, attributed to flash melting, results in a coefficient of friction of about 0.4, a value roughly consistent with flash melting, but perhaps slightly too high (Beeler and Tullis, 2003; Beeler et al., 2008; Goldsby and Tullis, 2003, 2011). The subsequent increase in friction to a value of the coefficient of friction of about 1.0 and the eventual decrease to values of about 0.6 do not appear to demonstrate significant weakening. However, expressing this steady-state shear resistance as a coefficient of friction is misleading. The data shown in **Figure 9** were collected at a normal stress of only 1.5 MPa, so the shear resistance is only 0.9 MPa. If this resistance is due to viscous shearing and if, as expected, the viscosity shows little or no dependence on pressure, then the shear resistance at higher pressures of perhaps 180 MPa at a depth of 10 km in the Earth might also be only 0.9 MPa. If so, the apparent coefficient of friction would be only 0.005 and the weakening would be extreme. Experiments conducted in Shimamoto's lab to somewhat higher normal stresses (up to several tens of MPa) seem to confirm this expectation, that is, the shear resistance once steady state is approached depends only weakly on normal stress (Di Toro et al., 2006). This is also expected from theoretical considerations (Fialko and Khazan, 2005). Although difficult to design and conduct, experiments at elevated confining pressure in which the melt cannot escape from the fault surface would be very valuable to explore further the weakening caused by macroscopic shear melting.

The presence of frozen friction melts (pseudotachylites) demonstrates that melting occurs during some earthquakes although they may only occur on low-displacement faults (Sibson and Toy, 2006a,b). Some investigators believe that pseudotachylites are much less abundant than expected, leading to speculation about the cause of this apparent paucity. Expectations of ubiquitous pseudotachylite are based on the assumption that earthquakes slip at velocities of about 1 m s^{-1} and occur at depth of a few kilometers with coefficients of friction of ~0.6, resulting in copious amounts of frictionally

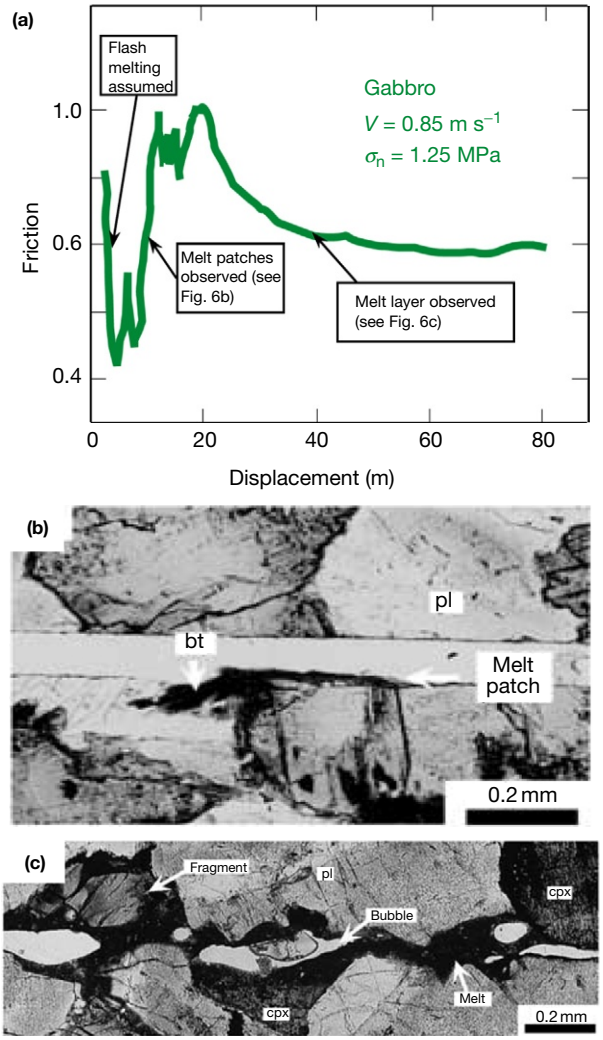


Figure 9 Results from friction experiments of Hirose and Shimamoto (2005) (their Figures 2a, 5a, and 5e). (a) Friction versus displacement, with initial weakening assumed due to flash melting, subsequent strengthening assumed due to increases in contact area as melt patches illustrated in (b) grow, and final weakening as melt layer illustrated in (c) thickens. (a) Redrawn from Hirose T and Shimamoto T (2005) Fractal dimension of molten surfaces as a possible parameter to infer the slip-weakening distance of faults from natural pseudotachylites. *Journal of Structural Geology* 25: 1569–1574.

generated heat (e.g., Kanamori and Heaton, 2000). Of course, one possibility is that one or more of the other weakening mechanisms discussed in this chapter reduce the frictional resistance during earthquakes to such low values that insufficient heat for melting is generated on most faults.

It is also very possible that pseudotachylites are often generated, but are either seldom preserved or simply seldom recognized. Some characteristics, especially the presence of glass or clearly discernible devitrification textures, are not likely to persist over geologic timescales, since recrystallization or deformation events can obliterate such evidence. Recrystallization should not eliminate the macroscopic morphology of pseudotachylite veins, but it might obscure the presence of layers of pseudotachylite on their generation surfaces unless the layers are thicker than typically observed. Deformation certainly has a good chance of destroying evidence of pseudotachylite in a fault zone, since subsequent shearing in the fault zone is to be expected. On the other hand, a layer of pseudotachylite can weld a fault surface together, making it stronger than before. Consequently, subsequent deformation may occur elsewhere, especially in an immature fault zone, preserving the earlier-formed pseudotachylite adjacent to more recently active fault strands. Pseudotachylite may be more common than is typically reported, but seldom recognized, since in many rocks, they are easy to overlook. Black pseudotachylite in white granite can be easily spotted, but in many dark rocks, they can be hard to discern unless the exposures are good and the observer is experienced. They are more likely to be generated in fault zones where the normal stress and hence the shear stress is high. They might also be expected to be more common in major fault zones where the slip displacements are high. Thus, they may be more common in major fault zones at depths of 5–10 km than in shallower faults. Unfortunately, there are not many exposures of major faults exhumed to this depth, especially those that have not been overprinted by later shallower slip.

Suffice it to say, the presence of pseudotachylites on some faults shows that shear melting sometimes occurs during earthquakes and the available experimental evidence at low normal stress shows that interfacial lubrication by frictional melt can be an efficient weakening mechanism. In addition, energy balance calculations in natural pseudotachylites based on thickness versus displacement relationships (Di Toro et al., 2006; Sibson, 1975) suggest that shear stress could be low in the presence of seismic melts.

This is the only high-speed weakening mechanism that operates with certainty during at least some earthquakes and it warrants continued study.

4.06.2.9 Interfacial Lubrication by Thixotropic Silica Gel

This weakening mechanism was discovered in the laboratory by Goldsby and Tullis (1998, 2002) and further investigated by Titone et al. (2001), Di Toro et al. (2004), and Hayashi and Tsutsumi (2010). A paper describing the role of water and the resulting textures is in preparation (Goldsby and Tullis, 2014). Extraordinary weakening that could not be explained by any other known mechanism was observed in a variety of experiments at relatively rapid slip rates and earthquake-like sliding displacements, eventually leading to the postulation of this mechanism. The mechanism involves the generation of a layer of silica gel on a fault surface by the interaction of silica in the sheared fault interface with water from the environment, either liquid water in the pore space or in fluid inclusions of the rock or water from the atmospheric humidity of the test environment. Silica gel is a noncrystalline, nonstoichiometric mixture of silica and water (Iler, 1979). The weakening observed in the experiments only occurs when the slip rate is high, above 1 mm s^{-1}, and the slip displacement is large, greater than several hundred millimeters (**Figure 10**). The steady-state friction coefficient drops from ~0.6 to ~0.25 as the slip velocity increases from 1 to 100 mm s^{-1} as shown in **Figure 11**. Following cessation of slip at speeds above 1 mm s^{-1}, the friction coefficient recovers to ~0.6 as a function

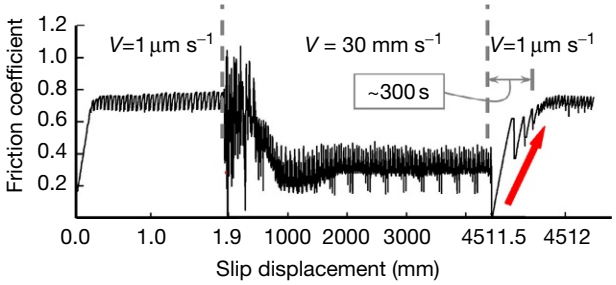


Figure 10 Friction versus displacement at low speed and high speed. The oscillations at 1 μm s^{-1} are due to stick–slip behavior, whereas those at 30 mm s^{-1} are artifacts due to rectification of frequent reversals of sliding direction. This high speed weakening due to gel lubrication requires a slip displacement of 0.5–1 m, whereas strength recovery at either low or zero velocity occurs as a function of time, not slip displacement. Reproduced from Figure 1a of Di Toro G, Goldsby DL, and Tullis TE (2004) Friction falls toward zero in quartz rock as slip velocity approaches seismic rates. *Nature* 427: 436–439.

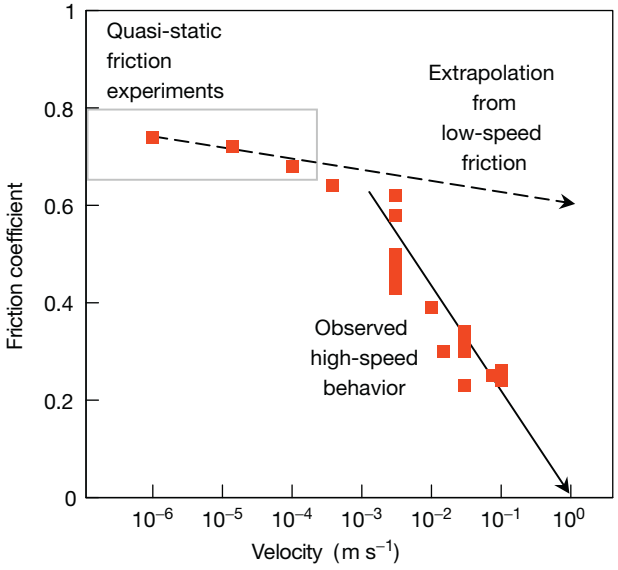


Figure 11 Friction versus slip velocity. The dramatic weakening above 1 mm s^{-1} is due to the formation of a thixotropic layer of silica gel, illustrated in **Figure 12**. The large reduction in resistance from gel weakening is in contrast to the extrapolated weakening expected from rate and state friction at low speed. Gray area shows typical seismic slip rates. Modified from Figure 3 of Di Toro G, Goldsby DL, and Tullis TE (2004) Friction falls toward zero in quartz rock as slip velocity approaches seismic rates. *Nature* 427: 436–439.

of time over a time interval from 300 s (**Figure 10**) to a few 1000 s, depending on the experimental details. As illustrated in **Figure 12**, SEM examination of the sliding surface shows a layer about 3 μm thick that was clearly soft and flowed during the experiments but is brittle after the experiments. These observations can be explained if the silica gel behaves as a thixotropic material (Shaw, 1980), a behavior that has been documented independently for silica gel (Kalinin et al., 2002; Serdobintseva and Kalinin, 2001).

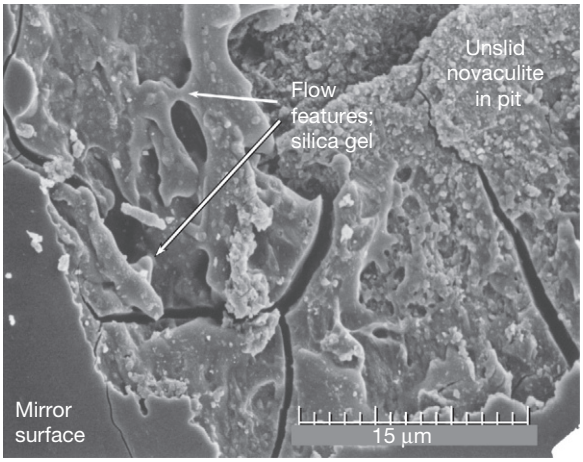


Figure 12 SEM micrograph of a portion of surface of novaculite sample slid 63 m at 3 mm s^{-1} in rotary shear experiments similar to those of Di Toro et al. (2004) and Goldsby and Tullis (2002). The field of view is at the edge of one of the many small pits in the surface that result from porosity in the sample. Silica gel has been smoothed over most of the surface, and at the edges of the pits, it shows flow features as it is smeared over the edge of the pit. The illustrated features show the gel had the ability to flow, but the fractures show that subsequently, its viscosity greatly increased.

The reason that silica gel can be weak when it experiences large strain at a high rate and strong when it is deformed slowly is that its strength depends on a competition between strain-dependent weakening and time-dependent strengthening (Ferguson and Kemblowski, 1991). Bonding between adjoining SiO_2 entities is disrupted by strain and this may be aided by the presence of water that can hydrolyze Si–O bonds. With no strain, the bonding is strengthened in a time-dependent way, perhaps accompanied by a rearrangement of the water. In the silica literature (e.g., Iler, 1979), the SiO_2 entities are described as being particles, similar to the micrometer- to submicrometer-sized particles that make up the regular structure of opal (Graetsch, 1994; Serdobintseva et al., 1998). However, it also seems possible that the SiO_2 entities also could be individual SiO_4^{-4} tetrahedrons or groups of tetrahedrons (Khavryuchenko and Sheka, 1994) and that the differences between the structure of the rapidly deformed silica gel and the slowly deformed silica gel might be the degree of polymerization, similar to the difference between the structure of a wet and a dry granite melt (Baker, 1998; Burnham, 1979; Shaw, 1972).

Weakening due to silica gel clearly depends on the presence of SiO_2 in appreciable amounts in the rock in the fault zone. Experiments employing rapid slip rates and large slips show that the weakening occurs, not only for pure quartz rock but also for silicate rocks with less silica. Gabbro, having only ~50 wt% SiO_2, does not show the weakening behavior, but granite and an albite feldspar rock, both with ~68 wt% SiO_2, reveal weakening (Roig Silva et al., 2004a,b). It is also worth noting that many rocks in nature having low percentages of SiO_2 nevertheless are cut by faults that contain abundant deformed quartz, introduced by hydrothermal vein deposition. An example is the fault zone of the Dixie Valley fault in Nevada (Power and Tullis, 1989). This fault zone has

experienced large historical earthquakes, in particular the M7.1 Fairview Peak and M6.8 earthquakes of 1954 (Caskey et al., 1996) and the M7.6 Pleasant Valley earthquake of 1915 (Doser, 1988). At the well-exposed 'mirror locality,' north of where the 1954 earthquakes occurred, the fault shows abundant quartz mineralization, even though the country rock is a gabbro (Power and Tullis, 1989). Thus, dramatic weakening during earthquakes may still occur during faulting in low-silica rocks due to the presence of mineralized quartz.

This weakening mechanism could be important for earthquakes. It is possible that at elevated temperature, the transition to thixotropic weakening might be shifted to velocities above 1 m s^{-1}, since the kinetics of gel dehydration might be enhanced. On the other hand, in hydrothermal environments as may be common on faults, elevated temperature in the presence of pressurized H_2O might promote gel formation. These speculations can be investigated by experiment. Verifying the operation of this weakening mechanism on faults may be difficult, because silica gel is even less likely to leave diagnostic traces than shear melting. The gel may lose its water and crystallize to chalcedony or other fine-grained quartz rocks. Very fine-grained deformed quartz layers in fault zones such as Dixie Valley (Power and Tullis, 1989) may plausibly have formed as a result of crystallization of silica gel, but proving this may not be possible. However, several authors do present evidence that silica gel existed within their fault zones (Faber et al., 2009, 2014; Kirkpatrick et al., 2013; Stel, 1981; Stel and Lankreyer, 1994), so it appears that this weakening mechanism may indeed operate during earthquake slip.

4.06.2.10 Rapid Superplastic Deformation of Fine-Grained Material

Fine-grained material is typically generated by comminution during frictional slip, whether at low or high slip velocities. Natural faults typically exhibit fine-grained material termed fault gouge within or adjacent to localized slip surfaces. The grain size of this material can range from nanoparticles (i.e., on the order of one to tens of nanometers in diameter) to tens of millimeters or larger, although it is termed fault breccia if the fragments are large. Recently, attention has been focused on the presence of nanoparticles in high-speed friction experiments and the term 'nanoparticle lubrication' or 'powder lubrication' has been used to describe the weakening that is observed to accompany the presence of the nanoparticles in these experiments (De Paola et al., 2011; Han et al., 2010, 2011; Hirose and Bystricky, 2007; Reches and Lockner, 2010). However, the presence of nanoparticles by themselves, even reduction to such small grain sizes that they appear amorphous to electron diffraction (Yund et al., 1990), is not sufficient to make the shear resistance of a fault low. Consequently, it is important to determine whether some new weakening mechanism is activated by the combination of very fine grain size and high strain rates. Rolling of particles might occur, but it is not apparent why this should occur at high, but not at low, rates. In some engineering situations, 'powder lubrication' does exist. One situation where powder lubrication occurs is in the presence of very weak layered materials, such as graphite, MoS_2, and WS_2 (Wornyoh et al., 2007), especially if they are nanoparticles in a fullerene-like configuration (Rapoport et al., 2003). The other situation, termed granular lubrication by Wornyoh et al. (2007), occurs if the pressure is low enough that the particles exist in some fluidized state where kinetic energy from particle impacts could explain the difference between the shear resistances at high and low rates (da Cruz et al., 2005). The latter situation may be thought of as similar to the acoustic fluidization mechanism discussed earlier and so the same concerns about its applicability to earthquake slip at elevated normal stress exist.

One intriguing mechanism that might explain observations of weakening in the presence of nanoparticles sheared at high slip rates has been offered by Green et al. (2010, 2014). The deformation mechanism is structural superplasticity in which the strain is primarily produced by grain boundary sliding, with either dislocation motion within the grains or diffusion providing the necessary accommodating deformation of the grains to allow them to slide past one another without creating gaps and overlaps (Langdon, 1994; Raj and Ashby, 1971). Typically, this is thought of as a slow deformation mechanism because of the need for diffusion to occur, but it can apparently occur at high strain rates (Chokshi et al., 1993; Padmanabhan and Basariya, 2009). The suggested explanation for superplasticity to occur at high strain rates is that if the grain size is on the order of nanometers, the diffusion distances are very small, so little time is needed. The rapid rate of heat input with rapid sliding and the consequent rise of temperature would contribute to the operation of this mechanism. This mechanism has also been proposed (Burnley and Green, 1991; Green and Burnley, 1989; Green and Houston, 1995; Schubnel et al., 2013) to explain both rapid slip in high-pressure experiments and the occurrence of deep earthquakes where normal stresses should be too high to allow conventional frictional slip. In this case, the nanoparticles are not produced by brittle comminution, but are fine-grained high-pressure phases produced when minerals stable at low pressure transform into ones stable at high pressure. In a somewhat related study, Kelemen and Hirth (2007) presented a model for intermediate-depth earthquakes due to accelerating creep rates and shear heating in which the deformation occurs by grain boundary sliding in fine-grained material, making the assumption that such fine-grained zones preexist.

Although this idea has only recently been proposed to explain the weakening in experiments that contain ultrafine-grained material and much more experimental work combined with detailed characterization of the sample material after the experiments, this mechanism could well be important for weakening during earthquakes.

Another possible explanation for the weakening of submicron powders sheared at high velocities as observed by Reches and Lockner (2010) has been presented by Sammis et al. (2011). This mechanism involves the behavior of a nanometer-thick layer of water surrounding the particles. At the longer particle-to-particle contact times associated with slow slip velocity, the water can be squeezed out of the contacts and the particle interactions result in high strength, whereas at short contact times associated with rapid slip velocities, the viscosity of the water layer prevents it from being removed from the contacts and the grains can more easily slip past one another lubricated by the water. They noted that at rapid enough shearing velocities that shear heating raises the temperature above 100 °C, the water layer will be removed and

weakening should cease. This would seem to eliminate this mechanism as an explanation for the many experiments involving nanoparticles that show low friction at sufficiently rapid rates that substantial temperatures are generated (Di Toro et al., 2011). It is not clear whether this mechanism would be expected to operate on water-saturated samples and on nominally dry samples involving only adsorbed water, but it seems likely that the arguments involving the role of time in squeezing water from particle to particle contacts would apply to saturated rocks as well, although the quantitative calculations in Sammis et al. (2011) would presumably require modification. This mechanism might apply even at elevated temperature in the water-saturated Earth's crust due to elevated water pressure.

4.06.3 Friction Resulting from High-Speed Weakening Mechanisms

For many of these weakening mechanisms, we do not know to what value the shear resistance might be reduced during coseismic slip, whereas for others, we have estimates of how low the shear stress might become. Recognizing that there is still uncertainty in the friction estimates and in whether these weakening mechanisms can operate during earthquakes, **Table 1** presents our current understanding of the effective friction coefficient possible from these mechanisms. Values are being presented to only one decimal point since our current knowledge does not warrant greater precision. These estimates are presented as effective coefficients of friction for reasons of convenience of comparison between mechanisms and for comparison with the commonly used friction coefficient for rock of ~0.6. A friction coefficient is not strictly an appropriate representation for some of the mechanisms. Examples include interfacial lubrication by frictional melt and superplastic deformation of fine-grained material, for both of which the resistance is better understood as a viscous resistance and any mechanism that involves dynamic reductions of the effective normal stress. In such cases, the effective friction coefficients should be considered to be the shear stress that resists dynamic sliding, divided by the effective normal stress that is expected to operate across a fault at a depth of about 10 km.

Table 1 Plausible values of shear resistance for proposed dynamic weakening mechanisms

Weakening mechanism	*Effective coefficient of friction* (τ/σ_n^{eff})
Dynamic normal stress reduction or loss of contact from normal interface vibrations	0.0–0.6
Dynamic normal stress reduction from elastic or permeability mismatch	?
Acoustic fluidization	?
Elastohydrodynamic lubrication	?
Thermal pressurization of pore fluid	0.0–0.6
Mineral breakdown-induced pressurization of pore fluid	0.0–0.6
'Flash' heating/melting at asperity contacts	~0.2
Interface lubrication by friction melt	0.0–0.6
Interface lubrication by thixotropic silica gel	<0.2
Rapid superplastic deformation of fine-grained material	<0.2

?, unknown.

4.06.4 Implications of Low Dynamic Friction for Earthquake Stress Drops and for Orientations and Magnitudes of Tectonic Stress

A complete discussion of the implications of low dynamic friction during earthquakes due to the various weakening mechanisms presented in **Table 1** is beyond the scope of this chapter, but some discussion is warranted. Low dynamic friction can result in slip occurring as self-healing pulses (Heaton, 1990) rather than in a more conventional crack-like manner (Rosakis, 2002). Some particularly important implications exist for the stress drops of earthquakes and/or for the levels of tectonic stresses typical of the Earth's crust and their orientations.

If the differential stress (the greatest principal compressive stress magnitude minus the least principal compressive stress magnitude) in the crust typically is 'high,' that is, approaching the differential stress required for frictional slip on optimally oriented faults with a static coefficient of friction of 0.6, as in situ stress measurements suggest (Hickman, 1991), and if shear stresses on faults during coseismic slip are essentially zero as **Table 1** suggests, then dynamic stress drops would be in the range of about 50–100 MPa for strike-slip earthquakes, for example, at depths of 5–10 km, respectively. However, stress drops are in the 1–10 MPa range for large earthquakes (Kanamori and Allen, 1986; Kanamori and Anderson, 1975).

Consequently, it seems possible that one or more of the three observations referred to in the previously mentioned paragraph, namely, high in situ measured differential stress, low dynamic friction (i.e., shear stress divided by effective normal stress, τ/σ_n), and low stress drops, are in error. It is certainly possible that earthquake slip does not occur at shear stresses approaching zero, and much more work on the weakening mechanisms that lead to this suggestion is unquestionably needed. On the other hand, given the number of potential weakening mechanisms that may reduce dynamic friction to nearly zero, it would not be surprising if one or more of them commonly operated during earthquakes. If this were the case, how can the dilemma presented by the combination of high in situ differential stress measurements, low dynamic friction, and estimates of low earthquake stress drops be resolved?

One point to make is that high shear stresses (i.e., half of the differential stress) do not need to exist everywhere in order to overcome high static frictional strength as an earthquake rupture propagates. (Here, we will refer to the static friction or static frictional strength as the value of the friction coefficient – i.e., the shear stress on the fault plane at which slip begins, divided by the effective normal stress on the fault plane – which must be overcome to initiate significant slip. The value of this friction coefficient is typically in the range of 0.6 to 0.8. It is typically only somewhat higher (~0.1) than the steady-state friction coefficient for sliding at slip velocities less than about 1 mm s^{-1} (see this velocity range in **Figure 11** for representative values of friction coefficient as a function of velocity). This static friction might be regarded as the unlabeled peak of

the solid curve in Figure 3 in Chapter 4.03) Only in the nucleation area of the earthquake is it necessary that static friction be overcome quasi-statically, possibly by having locally high shear stresses, low static friction, or low effective normal stress. Once the rupture begins to propagate, static friction can be overcome by the dynamic stress concentration at the tip of the propagating rupture (Lapusta and Rice, 2004), allowing a crust with relatively low average stress values to experience earthquakes with low stress drops and high static friction (see discussion following Eqn [39] in Chapter 4.03).

The difficulty with this model would seem to be the relatively high measured values of in situ differential stress. However, a possible resolution of this dilemma would be for the orientations of the principal stresses to be such that they are always at high angles to faults on which the stress drops are low and on which the dynamic friction is low. In this case, although the magnitudes of the differential stresses can be high and consistent with in situ stress measurements, the shear stresses on the fault would be low, permitting both low stress drops and low dynamic friction. This relationship between fault and stress orientations is of course one reason we believe the San Andreas Fault is weak (Mount and Suppe, 1987; Rice, 1992; Zoback et al., 1987). This relationship might only apply to major mature faults, and in fact, there is a tendency for higher stress drops to occur on intraplate earthquakes and those with longer repeat times than on interplate earthquakes (Kanamori and Allen, 1986; Kanamori and Anderson, 1975). Thus, for some small to moderate earthquakes on immature faults, observed stress drops do obtain values in the 30–60 MPa range (Chung and Gao, 1997; Dewberry and Crosson, 1996; Fukuyama et al., 1991; Huang and Yeh, 1999; Ide and Takeo, 1996; Kanamori et al., 1992; Ma and Kanamori, 1994; Mori, 1996; Takeo et al., 1993), that is, values as large as we would expect for high in situ differential stresses and low dynamic friction. Although the nearly orthogonal orientation relationship between major faults and principal stresses required by this model might seem to represent an infrequent situation, this in fact may be what one would expect to develop if faults are weak when they slip (i.e., have low dynamic friction), because the faults relax the shear stresses parallel to them, but not shear stress on other planes (Tullis, 2014).

Thus, an interesting possibility is that faults with low stress drops (i.e., in the 1–3 MPa range) tend to be the well-developed long-lived ones that have been able to relax the shear stresses parallel to them, whereas those with high stress drops (i.e., in the 10–60 MPa range) are less mature and have not yet slipped enough over many earthquake cycles to relax the shear stress on them and so tend to have the maximum and minimum principal stress orientations closer to 45° to them. This opens the possibility that for some earthquakes on appropriately oriented faults, the stress drops and the resulting ground motions might be substantially higher than is typically the case.

4.06.5 Conclusions

This brief review of our current understanding of high-speed weakening mechanisms that may result in low resistance to slip during earthquakes shows that many such mechanisms potentially exist. It is of course possible that further theoretical and experimental studies will discover new dynamic weakening mechanisms. It is also clear that much more remains to be understood about all of these weakening mechanisms. Although some of these mechanisms have been described in the literature for years, it is only recently that they have received renewed attention, so there is hope that it may not be too long before much more is known and this review will be out of date.

At present, it appears that many of these mechanisms may result in quite low shear resistance during coseismic slip (Table 1). The resulting low dynamic stress can be compatible with low stress drops and high magnitudes of in situ stress if principal stresses are oriented at high angles to faults with low stress drops.

Acknowledgments

The 2007 edition of this work was supported by USGS grants 05-HQGR0087 and 06-HQGR0068 and by the Southern California Earthquake Center (SCEC). The 2014 revised edition was also supported by the SCEC, which is funded by NSF Cooperative Agreement EAR-1033462 and USGS Cooperative Agreement G12AC20038. The SCEC contribution number for this chapter is 1916. I thank David Goldsby, Giulio Di Toro, Nick Beeler, Jim Rice, Toshi Shimamoto, Greg Hirth, and Tom Heaton for their many helpful discussions on these topics; Jim Brune and Rasool Anooshehpoor for checking the section on normal interface vibrations; Jay Melosh for an update on acoustic fluidization; David Goldsby for providing Figures 4–6 and 12; Vikas Prakash for providing Figures 7 and 8; and Giulio Di Toro for providing the materials for Figure 9. I would especially like to thank David Goldsby and Giulio Di Toro, Hiroo Kanamori, and Nick Beeler for their helpful reviews that improved the presentation and the science content of the 2007 version of this manuscript and Nick Beeler for his very helpful discussions about the work since 2007. I thank an anonymous reviewer for making several helpful suggestions, including pointing out several relevant recent papers I had overlooked that needed inclusion and discussion.

References

Adams GG (1995) Self-excited oscillations of two elastic half-spaces sliding with constant coefficient of friction. *Journal of Applied Mechanics* 62: 867–872.

Adams GG (1998) Steady sliding of two elastic half-spaces with friction reduction due to interface stick-slip. *Journal of Applied Mechanics* 65: 470–475.

Adams GG (2001) An intersonic slip pulse at a frictional interface between dissimilar solids. *Journal of Applied Mechanics* 68: 81–86.

Andrews DJ (1976) Rupture velocity of plane strain shear cracks. *Journal of Geophysical Research* 81: 5679–5687.

Andrews DJ (2002) A fault constitutive relation accounting for thermal pressurization of pore fluid. *Journal of Geophysical Research* 107: 12.

Andrews DJ and Ben-Zion Y (1997) Wrinkle-like slip pulse on a fault between different materials. *Journal of Geophysical Research* 102: 553–571.

Andrews DJ and Harris RA (2005) The wrinkle-like slip pulse is not important in earthquake dynamics. *Geophysical Research Letters* 32: L23303. http://dx.doi.org/10.1029/2005GL023996.

Andrews DJ and Harris RA (2006) Reply to comment by Y. Ben-Zion on "The wrinkle-like slip pulse is not important in earthquake dynamics". *Geophysical Research Letters* 33(6): L06311.

Anooshehpoor A and Brune JN (1994a) Dynamics and frictional heat generation in a foam rubber model of earthquake stick-slip. In: *Proceedings of Workshop LXIII The Mechanical Involvement of Fluids in Faulting, June 6–10, 1993, Fish Camp, CA*, 443–459, pp. USGS Red Book NEHRP Open-File Report 94-228.

Anooshehpoor A and Brune JN (1994b) Frictional heat generation and seismic radiation in a foam rubber model of earthquakes. *Pure and Applied Geophysics* 142: 735–747.

Anooshehpoor A and Brune JN (1999) Wrinkle-like Weertman pulse at the interface between two blocks of foam rubber with different velocities. *Geophysical Research Letters* 26(13): 2025–2028.

Archard JF (1958) The temperature of rubbing surfaces. *Wear* 2: 438–455.

Baker DR (1998) Granitic melt viscosity and dike formation. *Journal of Structural Geology* 20(9–10): 1395–1404.

Beeler NM and Tullis TE (1997) The roles of time and displacement in velocity-dependent volumetric strain of faults. *Journal of Geophysical Research* 102: 22595–22609.

Beeler NM and Tullis TE (2003) Constitutive relationships for fault strength due to flash-heating. In: *Paper Presented at 2003 SCEC Annual Meeting Proceedings and Abstracts, Southern California Earthquake Center, Palm Springs, California*.

Beeler NM, Tullis TE, Blanpied ML, and Weeks JD (1996) Frictional behavior of large displacement experimental faults. *Journal of Geophysical Research* 101: 8697–8715.

Beeler NM, Tullis TE, and Goldsby DL (2008) Constitutive relationships and physical basis of fault strength due to flash-heating. *Journal of Geophysical Research* 113: B01401, http://dx.doi.org/01410.01029/02007JB004988.

Beeler NM, Tullis TE, and Goldsby DL (2008) Constitutive relationships and physical basis of fault strength due to flash-heating. *Journal of Geophysical Research* 113: B01401. http://dx.doi.org/10.01029/02007JB004988.

Ben-Zion Y (2001) Dynamic ruptures in recent models of earthquake faults. *Journal of the Mechanics and Physics of Solids* 49(9): 2209–2244.

Ben-Zion Y (2006) Comment on "The wrinkle-like slip pulse is not important in earthquake dynamics" by D. J. Andrews and R. A. Harris. *Geophysical Research Letters* 33: L06310. http://dx.doi.org/10.01029/02005GL025372.

Ben-Zion Y and Andrews DJ (1998) Properties and implication of dynamic rupture along a material interface. *Bulletin of the Seismological Society of America* 88: 1085–1094.

Ben-Zion Y and Huang Y (2002) Dynamic rupture on an interface between a compliant fault zone layer and a stiffer surrounding solid. *Journal of Geophysical Research* 107(B2): 2042. http://dx.doi.org/10.1029/2001JB000254.

Ben-Zion Y and Shi Z (2005) Dynamic rupture on a material interface with spontaneous generation of plastic strain in the bulk. *Earth and Planetary Science Letters* 236: 486–496. http://dx.doi.org/10.1016/j.epsl.2005.1003.1025.

Bjornerud A and Magloughlin JF (2004) Pressure-related feedback processes in the generation of pseudotachylytes. *Journal of Structural Geology* 26(12): 2317–2323.

Blok H (1937) Measurements of temperature flashes on gear teeth under extreme pressure. *Proceedings of the Institution of Mechanical Engineers* 1-2: 14–20.

Bowden FB and Persson PA (1960) Deformation heating and melting of solids in high speed friction. *Proceedings of the Royal Society of London, Series A* 260: 433–458.

Bowden FP and Ridler KEW (1936) Physical properties of surfaces. III. The surface temperature of sliding metals: The temperature of lubricated surfaces. *Proceedings of the Royal Society of London, Series A* 154: 640–656.

Bowden FP, Stone MA, and Tudor GK (1947) Hot spots on rubbing surfaces and the detonation of explosives by friction. *Proceedings of the Royal Society of London, Series A* 188: 329–349.

Bowden FP and Thomas PH (1954) The surface temperature of sliding solids. *Proceedings of the Royal Society of London, Series A* 223: 29–40.

Brantut N, Han R, Shimamoto T, Findling N, and Schubnel A (2011) Fast slip with inhibited temperature rise due to mineral dehydration: Evidence from experiments on gypsum. *Geology* 39: 59–62.

Brantut N, Schubnel A, Corvisier J, and Sarout J (2010) Thermochemical pressurization of faults during coseismic slip. *Journal of Geophysical Research* 115: B05314. http://dx.doi.org/10.01029/02009JB006533.

Brantut N, Schubnel A, Rouzaud J-N, Brunet F, and Shimamoto T (2008) High-velocity frictional properties of a clay-bearing fault gouge and implications for earthquake mechanics. *Journal of Geophysical Research* 113: B10401. http://dx.doi.org/10.11029/12007JB005551.

Brodsky EE and Kanamori H (2001) Elastohydrodynamic lubrication of faults. *Journal of Geophysical Research* 106: 16357–16374.

Brown SR, Olsson WA, and Brune JN (1991) Experimental observation of interface separation waves during stick-slip. *EOS, Transactions American Geophysical Union* 72: 326.

Brune JN (1991) Seismic source dynamics, radiation and stress. *Reviews of Geophysics, Supplement* IUGG 1987–1999: 688–699.

Brune JN (1996) Particle motions in a physical model of shallow angle thrust faulting. *Proceedings of the Indian Academy of Science* 105(2): L197–L206.

Brune J (2001) Fault-normal dynamic loading and unloading: An explanation for 'nongouge' rock powder and lack of fault-parallel shear bands along the San Andreas Fault. *EOS, Transactions American Geophysical Union* 82(47), Fall Meeting Supplement.

Brune JN, Brown S, and Johnson PA (1993) Rupture mechanism and interface separation in foam rubber models of earthquakes: A possible solution to the heat flow paradox and the paradox of large overthrusts. *Tectonophysics* 218: 59–67.

Brune JN, Johnson PA, and Slater C (1990) Nucleation, predictability, and rupture mechanism in foam rubber models of earthquakes. *Journal of Himalayan Geology* 1: 155–166.

Brune JN and Thatcher W (2002) Strength and energetics of active fault zones. *International Handbook of Earthquake and Engineering Seismology*, vol. 81A, pp. 569–588, Academic Press.

Burnham CW (1979) The importance of volatile constituents. In: Yoder HS Jr. (ed.) *The Evolution of Igneous Rocks: Fiftieth Anniversary Perspectives*, pp. 439–482. Princeton, NJ: Princeton University Press.

Burnley PC and Green HW II (1991) Faulting associated with the olivine to spinel transformation in Mg_2GeO_4 and its implications for deep-focus Earthquakes. *Journal of Geophysical Research* 96(B1): 425–443.

Byerlee JD (1978) Friction of rocks. *Pure and Applied Geophysics* 116: 615–626.

Caskey SJ, Wesnousky SG, Zhang P, and Slemmons DB (1996) Surface faulting of the 1954 Fairview Peak (MS 7.2) and Dixie Valley (MS 6.8) earthquakes, central Nevada. *Bulletin of the Seismological Society of America* 86(3): 761–787.

Castro RR, Anderson JG, and Brune JN (1991) Origin of high P/S spectral ratios from the Guerrero accelerograph array. *Bulletin of the Seismological Society of America* 81(6): 2268–2288.

Chester FM and Chester JS (1998) Ultracataclasite structure and friction processes of the Punchbowl fault San Andreas system, California. *Tectonophysics* 295: 199–221.

Chester FM, Chester JS, Kirschner DL, Schulz SE, and Evans JP (2004) Structure of large-displacement strike-slip fault zones in the brittle continental crust. In: Karner GD, Taylor B, Driscoll NW, and Kohlstedt DL (eds.) *Rheology and Deformation in the Lithosphere at Continental Margins*, pp. 223–260. New York: Columbia University Press.

Chester FM, Evans JP, and Biegel RL (1993) Internal structure and weakening mechanisms of the San Andreas Fault. *Journal of Geophysical Research* 98(B1): 771–786.

Chokshi AH, Mukherjee AK, and Langdon TG (1993) Superplasticity in advanced materials. *Materials Science and Engineering* R10: 237–274.

Chung WY and Gao H (1997) The Greenland earthquake of 11 July 1987 and postglacial fault reactivation along a passive margin. *Bulletin of the Seismological Society of America* 87(4): 1058–1068.

Cochard A and Rice J (2000) Fault rupture between dissimilar materials: Ill-posedness, regularization, and slip-pulse response. *Journal of Geophysical Research* 105: 25891–25907.

Collins GS and Melosh HJ (2003) Acoustic fluidization and the extraordinary mobility of sturzstroms. *Journal of Geophysical Research: Solid Earth* 108(B10).

Cottrell AH (1965) *Dislocations and Plastic Flow in Crystals*, p. 223. London: Oxford University Press.

da Cruz F, Emam S, Prochnow M, Roux JN, and Chevoir F (2005) Rheophysics of dense granular materials: Discrete simulation of plane shear flows. *Physical Review E* 72(2): 021309, http://dx.doi.org/10.1103/PhysRevE.72.021309.

De Paola N, Hirose T, Mitchell T, Di Toro G, Viti C, and Shimamoto T (2011) Fault lubrication and earthquake propagation in thermally unstable rocks. *Geology* 39(1): 35–38. http://dx.doi.org/10.1130/G31398.

Dewberry SR and Crosson RS (1996) The M(D) 5.0 earthquake of 29 January 1995 in the Puget lowland of western Washington: An event on the Seattle fault? *Bulletin of the Seismological Society of America* 86(4): 1167–1172.

Di Toro G, Goldsby DL, and Tullis TE (2004) Friction falls toward zero in quartz rock as slip velocity approaches seismic rates. *Nature* 427: 436–439.

Di Toro G, Han R, Hirose T, et al. (2011) Fault lubrication during earthquakes. *Nature* 471: 494–498. http://dx.doi.org/10.1038/nature09838.

Di Toro G, Hirose T, Nielsen S, Pennacchioni G, and Shimamoto T (2006) Natural and experimental evidence of melt lubrication of faults during earthquakes. *Science* 311(5761): 647–649.

Di Toro G, Nielsen S, and Pennacchioni G (2005a) Earthquake rupture dynamics frozen in exhumed ancient faults. *Nature* 436(7053): 1009–1012.

Di Toro G and Pennacchioni G (2004) Superheated friction-induced melts in zoned pseudotachylytes within the Adamello tonalites (Italian Southern Alps). *Journal of Structural Geology* 26(10): 1783–1801.

Di Toro G and Pennacchioni G (2005) Fault plane processes and mesoscopic structure of a strong-type seismogenic fault in tonalites (Adamello batholith, Southern Alps). *Tectonophysics* 402(1–4): 55–80.

Di Toro G, Pennacchioni G, and Teza G (2005b) Can pseudotachylytes be used to infer earthquake source parameters? An example of limitations in the study of exhumed faults. *Tectonophysics* 402(1–4): 3–20.

Dieterich JH (1972) Time-dependent friction in rocks. *Journal of Geophysical Research* 77: 3690–3697.

Dieterich JH (1978) Time-dependent friction and the mechanics of stick slip. *Pure and Applied Geophysics* 116: 790–806.

Dieterich JH (1979) Modeling of rock friction. 1: Experimental results and constitutive equations. *Journal of Geophysical Research* 84: 2161–2168.

Dieterich JH and Kilgore BD (1994) Direct observation of frictional contacts: New insights for state-dependent properties. *Pure and Applied Geophysics* 143: 283–302.

Dieterich JH and Kilgore BD (1996) Imaging surface contacts: Power law contact distributions and contact stresses in quartz, calcite, glass, and acrylic plastic. *Tectonophysics* 256: 219–239.

Dor O, Ben-Zion Y, Rockwell TK, and Brune JN (2006a) Pulverized rocks in the Mojave section of the San Andreas Fault zone. *Earth and Planetary Science Letters* 245: 642–654. http://dx.doi.org/10.1016/j.epsl.2006.1003.1034.

Dor O, Rockwell TK, and Ben-Zion Y (2006b) Geologic observations of damage asymmetry in the structure of the San Jacinto, San Andreas and Punchbowl faults in southern California: A possible indicator for preferred rupture propagation direction. *Pure and Applied Geophysics* 163 :301–349. http://dx.doi.org/10.1007/s00024-00005-00023-00029.

Doser DI (1988) Source parameters of earthquakes in the Nevada Seismic Zone, 1915–1943. *Journal of Geophysical Research* 93(B12): 15001–15015.

Dunham EM and Rice JR (2008) Earthquake slip between dissimilar poroelastic materials. *Journal of Geophysical Research: Solid Earth* 113: B09304. http://dx.doi.org/10.01029/02007JB005405.

Ettles CMM (1986) The thermal control of friction at high sliding speeds. *Journal of Tribology – Transactions of the ASME* 108(1): 98–104.

Faber C, Rowe CD, Miller JA, Backeberg N, and Sylvester FY (2009) Possible silica gel in the Olive Fault, Naukluft Nappe Complex, Namibia: A geologic record of dynamic weakening during continental orogenesis. In: *American Geophysical Union Fall Meeting*, Abstract T53C-1606.

Faber C, Rowe CD, Miller JA, Fagereng A, and Neethling JH (2014) Silica gel in a fault slip surface: field evidence for paleo-earthquakes?. *Journal of Structural Geology* 69: 108–121.

Faulkner DR, Mitchell TM, Behnsen J, Hirose T, and Shimamoto T (2011) Stuck in the mud? Earthquake nucleation and propagation through accretionary forearcs. *Geophysical Research Letters* 38: L18303. http://dx.doi.org/10.11029/12011GL048552.

Ferguson J and Kemblowski Z (1991) *Applied Fluid Rheology*, p. 323. New York: Elsevier.

Ferre EC, Allen JL, and Lin AM (2005) Pseudotachylytes and seismogenic friction: An introduction to current research. *Tectonophysics* 402(1–4): 1–2.

Fialko Y (2004) Temperature fields generated by the elastodynamic propagation of shear cracks in the Earth. *Journal of Geophysical Research: Solid Earth* 109(B1).

Fialko Y and Khazan Y (2005) Fusion by earthquake fault friction: Stick or slip? *Journal of Geophysical Research* 110: B12407. http://dx.doi.org/10.1029/2005JB003869.

Fukuyama E, Kinoshita S, and Yamamizu F (1991) Unusual high-stress drop subevent during the M5.5 earthquake, the largest event of the Ito-Oki swarm activity. *Geophysical Research Letters* 18(4): 641–644.

Goldsby DL, Passelegue F, Tullis TE, and Fabbri O (2011) The influence of ambient fault temperature on flash-heating phenomena. In: *AGU Fall Meeting Abstracts*, Abstract #T23G-07.

Goldsby DL and Tullis TE (1997) Shear heating induced pressurization of pore fluid as a dynamic fault weakening mechanism. *EOS, Transactions American Geophysical Union* 78: 472, Fall Meeting Supplement.

Goldsby DL and Tullis TE (1998) Experimental observations of frictional weakening during large and rapid slip. *EOS, Transactions American Geophysical Union* 97: F610.

Goldsby DL and Tullis TE (2002) Low frictional strength of quartz rocks at subseismic slip rates. *Geophysical Research Letters* 29(17): 1844. http://dx.doi.org/10.1029/2002GL015240.

Goldsby DL and Tullis TE (2003) Flash heating/melting phenomena for crustal rocks at (nearly) seismic slip rates. In: *Paper presented at 2003 SCEC Annual Meeting Proceedings and Abstracts, Southern California Earthquake Center, Palm Springs, California*.

Goldsby DL and Tullis TE (2007) In: *Proceedings and Abstracts 2007 SCEC Annual Meeting, Vol. 17* .

Goldsby DL and Tullis TE (2011) Flash heating leads to low strength of crustal rocks at earthquake slip rates. *Science* 334: 216–218. http://dx.doi.org/10.1126/science.1207902.

Goldsby DL and Tullis TE (2014) Silica gel weakening in high speed rock friction experiments. *Journal of Geophysical Research*, in preparation.

Goldschmidt VM (1943) Om friksjonsglass (pseudo-tachylitt) i fjellkjeden. *GFF* 65: 83–84.

Graetsch H (1994) Structural characteristics of opaline and microcrystalline silica minerals. In: Heaney PJ, Prewitt CT, and Gibbs GV (eds.) *Silica, Physical Behavior, Geochemistry, and Materials Applications*. *Reviews in Mineralogy*, vol. 29, pp. 209–232. Washington, DC: Mineralogical Society of America.

Green HW II and Burnley PC (1989) A new self-organizing mechanism for deep-focus earthquakes. *Nature* 341(6244): 733–737.

Green HW II and Houston H (1995) The mechanics of deep earthquakes. *Annual Review of Earth and Planetary Sciences* 23: 169–213.

Green HW, Lockner DA II, Bozhilov KN, Maddon A, Beeler NM, and Reches Z (2010) Nanometric gouge in high-speed shearing experiments: Superplasticity? In: *American Geophysical Union Fall Meeting, San Francisco*, Abstract #T31D-08.

Green HW II, Shi F, Bozhilov K, Xia G, and Reches Z (2014) A nanometric mechanism for earthquake sliding. *Nature Geoscience*, submitted.

Hamrock BJ (1994) *Fundamentals of Fluid Film Lubrication*. New York: McGraw-Hill.

Han R, Hirose T, and Shimamoto T (2010) Strong velocity weakening and powder lubrication of simulated carbonate faults at seismic slip rates. *Journal of Geophysical Research* 115: B03412. http://dx.doi.org/10.01029/02008JB006136.

Han R, Hirose T, Shimamoto T, Lee Y, and Ando J (2011) Granular nanoparticles lubricate faults during seismic slip. *Geology* 39: 599–602. http://dx.doi.org/10.1130/G31842.31841.

Han R, Shimamoto T, Hirose T, Ree J-H, and Ando J (2007) Ultralow friction of carbonate faults caused by thermal decomposition. *Science* 316: 878–881.

Harris RA and Day SM (1997) Effects of a low-velocity zone on a dynamic rupture. *Bulletin of the Seismological Society of America* 87(5): 1267–1280.

Harris RA and Day SM (2005) Material contrast does not predict earthquake rupture propagation direction. *Geophysical Research Letters* 32(23).

Hayashi N and Tsutsumi A (2010) Deformation textures and mechanical behavior of a hydrated amorphous silica formed along an experimentally produced fault in chert. *Geophysical Research Letters* 37: L12305. http://dx.doi.org/10.11029/12010GL042943.

Heaton TH (1990) Evidence for and implications of self-healing pulses of slip in earthquake rupture. *Physics of the Earth and Planetary Interiors* 64: 1–20.

Hickman SH (1991) Stress in the lithosphere and the strength of active faults, U.S. Nat. Rep. Int. Union of Geod. and Geophys. 1987–1990. *Reviews of Geophysics* 29: 759–775.

Hirose T (2002) *Experimental and Field Studies on the Frictional Melting Along Faults and Implications to the Earthquake Generation Processes*. PhD Thesis, Kyoto University, Kyoto, Japan.

Hirose T and Bystricky M (2007) Extreme dynamic weakening of faults during dehydration by coseismic shear heating. *Geophysical Research Letters* 34: L14311. http://dx.doi.org/10.11029/12007GL030049.

Hirose T and Shimamoto T (2003) Fractal dimension of molten surfaces as a possible parameter to infer the slip-weakening distance of faults from natural pseudotachylites. *Journal of Structural Geology* 25: 1569–1574.

Hirose T and Shimamoto T (2005) Growth of a molten zone as a mechanism of slip weakening of simulated faults in gabbro during frictional melting. *Journal of Geophysical Research* 110: B05202. http://dx.doi.org/10.01029/02004JB003207.

Hirose T, Tsutsumi A, and Shimamoto T (2000) High-velocity slip weakening on gabbro during frictional melting. *EOS, Transactions, American Geophysical Union* 81(48), Fall Meeting Supplement, Abstract T12B-09.

Huang WG and Yeh YT (1999) Characteristics of seismic source spectra from the Chia-Yi and Tai-Nan area of Taiwan. *Terrestrial, Atmospheric and Oceanic Sciences* 10(2): 415–446.

Hull D (1965) *Introduction to Dislocations*, p. 259. Oxford: Pergamon Press.

Ide S and Takeo M (1996) The dynamic rupture process of the 1993 Kushiro-oki earthquake. *Journal of Geophysical Research: Solid Earth* 101(B3): 5661–5675.

Iler RK (1979) *The Chemistry of Silica*, p. 720. New York: Wiley.

Jaegar JC (1942) Moving sources of heat and temperature at a sliding contact. *Journal and Proceedings of the Royal Society of New South Wales* 76: 203–224.

Jeffreys H (1942) On the mechanics of faulting. *Geological Magazine* 79: 291–295.

Johnson KL (1985) *Contact Mechanics*, p. 452. Cambridge: Cambridge University Press.

Junger JA and Tullis TE (2003) Fault roughness and matedness suggest significant fault-interface dilatancy with slip. *EOS, Transactions American Geophysical Union* 84(46), Fall Meeting Supplement, Abstract S51B-03.

Kalinin DV, Serdobintseva VV, and Kuznetsova YV (2002) Thixotropy in supramolecular crystallization and genesis of macrostructures of precious opal. *Geologiya I Geofizika* 43(11): 1002–1008.

Kanamori H and Allen CR (1986) Earthquake repeat times and average stress drop. In: Das S, Boatwright J, and Scholz CH (eds.) *Geophysical Monograph*, vol. 37. Washington, DC: American Geophysical Union.

Kanamori H and Anderson DL (1975) Theoretical basis of some empirical relations in seismology. *Bulletin of the Seismological Society of America* 65(5): 1073–1095.

Kanamori H and Heaton TH (2000) Microscopic and macroscopic physics of earthquakes. In: Rundle J, Turcotte DL, and Klein W (eds.) *GeoComplexity and the Physics of Earthquakes*, pp. 147–163. Washington, DC: American Geophysical Union.

Kanamori H, Thio HK, Dreger D, Hauksson E, and Heaton T (1992) Initial investigation of the Landers, California, earthquake of 28 June 1992 using TERRAscope. *Geophysical Research Letters* 19(22): 2267–2270.

Kelemen PB and Hirth G (2007) A periodic shear-heating mechanism for intermediate-depth earthquakes in the mantle. *Nature* 446: 787–790. http://dx.doi.org/10.1038/nature05717.

Khavryuchenko VD and Sheka EF (1994) Computational modeling of amorphous silica. 3. Modeling the initial structures – Silica-gel. *Journal of Structural Chemistry* 35(3): 299–304.

Kirkpatrick JD, Rowe CD, White JC, and Brodsky EE (2013) Silica gel formation during fault slip: Evidence from the rock record. *Geology* 41: 1015–1018. http://dx.doi.org/10.1130/G34483.34481.

Kitajima H, Chester FM, and Chester JS (2011) Dynamic weakening of gouge layers in high-speed shear experiments: Assessment of temperature-dependent friction, thermal pressurization, and flash heating. *Journal of Geophysical Research* 116: B08309. http://dx.doi.org/10.01029/02010JB007879.

Kitajima H, Chester JS, Chester FM, and Shimamoto T (2010) High-speed friction of disaggregated ultracataclasite in rotary shear: Characterization of frictional heating, mechanical behavior, and microstructure evolution. *Journal of Geophysical Research* 115: B08408. http://dx.doi.org/10.01029/02009JB007038.

Kohli AH, Goldsby DL, Hirth G, and Tullis TE (2011) Flash weakening of serpentinite at near-seismic slip rates. *Journal of Geophysical Research* 116: B03202. http://dx.doi.org/10.1029/2010JB007833.

Lachenbruch AH (1980) Frictional heating, fluid pressure, and the resistance to fault motion. *Journal of Geophysical Research* 85: 6249–6272.

Langdon TG (1994) A unified approach to grain boundary sliding in creep and superplasticity. *Acta Metallurgica* 42: 2437–2443. http://dx.doi.org/10.1016/0956-7151(2494)90322-90320.

Lapusta N and Rice JR (2004) Earthquake sequences on rate and state faults with strong dynamic weakening. *EOS, Transactions, American Geophysical Union* 85(47), Fall Meeting Supplement, T22A-05.

Lee TC and Delaney PT (1987) Frictional heating and pore pressure rise due to a fault slip. *Geophysical Journal of the Royal Astronomical Society* 88(3): 569–591.

Lim SC and Ashby MF (1987) Wear mechanism maps. *Acta Metallurgica* 35: 1–24.

Lim SC, Ashby MF, and Brunton JH (1989) The effects of sliding conditions on the dry friction of metals. *Acta Metallurgica* 37(3): 767–772.

Lin A and Shimamoto T (1998) Selective melting processes as inferred from experimentally generated pseudotachylytes. *Journal of Asian Earth Sciences* 16: 533–545.

Lockner D, Naka H, Tanaka H, Ito H, and Ikeda R (2000) Permeability and strength of core samples from the Nojima fault of the 1995 Kobe earthquake. In: Ito H, et al. (ed.). *Proceedings of the International Workshop on the Nojima Fault Core and Borehole Data Analysis, Tsukuba, Japan, November 22–23,* pp. 147–152, U.S. Geological Survey Open File Report, 00-129.

Ma KF and Kanamori H (1994) Broad-band wave-form observation of the 28 June 1991 Sierra-Madre earthquake sequence ($M(L) = 5.8$). *Bulletin of the Seismological Society of America* 84(6): 1725–1738.

Maddock R (1986) Frictional melting in landslide-generated frictionites (hyalomylonites) and fault-generated pseudotachylytes – Discussion. *Tectonophysics* 128: 151–153.

Magloughlin JF (1989) The nature and significance of pseudotachylite from the Nason Terrane, North Cascade Mountains, Washington. *Journal of Structural Geology* 11(7): 907–917.

Magloughlin JF (1992) Microstructural and chemical-changes associated with cataclasis and frictional melting at shallow crustal levels – The cataclasite pseudotachylyte connection. *Tectonophysics* 204(3–4): 243–260.

Magloughlin JF and Spray JG (1992) Frictional melting processes and products in geological materials – Introduction and discussion. *Tectonophysics* 204(3–4): 197–206.

Major JJ and Pierson TC (1992) Debris flow rheology – Experimental-analysis of fine-grained slurries. *Water Resources Research* 28(3): 841–857.

Marone CJ (1998) Laboratory-derived friction constitutive laws and their application to seismic faulting. *Annual Review of Earth and Planetary Sciences* 26: 643–696.

Marone CJ, Raleigh CB, and Scholz CH (1990) Frictional behavior and constitutive modeling of simulated fault gouge. *Journal of Geophysical Research* 95: 7007–7025.

Mase CW and Smith L (1985) Pore-fluid pressures and frictional heating on a fault surface. *Pure and Applied Geophysics* 122: 583–607.

Mase CW and Smith L (1987) Effects of frictional heating on the thermal hydrologic and mechanical response of a fault. *Journal of Geophysical Research* 92: 6249–6272.

McKenzie DP and Brune JN (1972) Melting of fault planes during large earthquakes. *Geophysical Journal of the Royal Astronomical Society* 29: 65–78.

Melosh HJ (1979) Acoustic fluidization: A new geologic process? *Journal of Geophysical Research* 84: 7513–7520.

Melosh HJ (1996) Dynamic weakening of faults by acoustic fluidization. *Nature* 379: 601–606.

Melosh HJ and Girdner KKETAGU (1995) Rheology of vibrated granular materials: Application to long runout landslides. *EOS, Transactions American Geophysical Union* 76(46): F270.

Melosh HJ and Ivanov BA (1999) Impact crater collapse. *Annual Review of Earth and Planetary Sciences* 27: 385–415.

Mizoguchi K, Hirose T, Shimamoto T, and Fukuyama E (2009) High-velocity frictional behavior and microstructure evolution of fault gouge obtained from Nojima fault, southwest Japan. *Tectonophysics* 471: 285–296. http://dx.doi.org/10.1016/J.Tecto.2009.1002.

Molinari A, Estrin Y, and Mercier S (1999) Dependence of the coefficient of friction on the sliding conditions in the high velocity range. *Journal of Tribology – Transactions of the ASME* 121(1): 35–41.

Mori J (1996) Rupture directivity and slip distribution of the M 4.3 foreshock to the Joshua Tree earthquake, Southern California. *Bulletin of the Seismological Society of America* 86(3): 805–810.

Morrow CA and Lockner DA (1997) Permeability and porosity of the Illinois UPH 3 drillhole granite and a comparison with other deep drillhole rocks. *Journal of Geophysical Research: Solid Earth* 102(B2): 3067–3075.

Morrow CA, Moore DE, and Lockner DA (2001) Permeability reduction in granite under hydrothermal conditions. *Journal of Geophysical Research: Solid Earth* 106(B12): 30551–30560.

Morrow CA, Shi LQ, and Byerlee JD (1984) Permeability of fault gouge under confining pressure and shear stress. *Journal of Geophysical Research* 99: 3193–3200.

Mount VS and Suppe J (1987) State of stress near the San Andreas Fault: Implications for wrench tectonics. *Geology* 15: 1143–1146.

Nielsen S, Di Toro G, Hirose T, and Shimamoto T (2008) Frictional melt and seismic slip. *Journal of Geophysical Research* 113: B01308. http://dx.doi.org/10.01029/02007JB005122.

Nielsen S, Mosca P, Giberti G, Di Toro G, Hirose T, and Shimamoto T (2010) On the transient behavior of frictional melt during seismic slip. *Journal of Geophysical Research* 115: B10301. http://dx.doi.org/10.11029/12009JB007020.

Niemeijer A, Di Toro G, Griffith WA, Bistacchi A, Smith SA, and Nielsen S (2012) Inferring earthquake physics and chemistry using an integrated field and laboratory approach. *Journal of Structural Geology* 39: 2–36. http://dx.doi.org/10.1016/j.jsg.2012.1002.1018.

Noda H and Shimamoto T (2005) Thermal pressurization and slip-weakening distance of a fault: An example of the Hanaore fault, southwest Japan. *Bulletin of the Seismological Society of America* 95(4): 1224–1233.

Padmanabhan KA and Basariya MR (2009) Mesoscopic grain boundary sliding as the rate controlling process for high strain rate superplastic deformation. *Materials Science and Engineering A* 527: 225–234.

Passelegue F, Goldsby DL, and Fabbri O (2014) The influence of ambient fault temperature on flash-heating phenomena. *Geophysical Research Letters* 41(3): 828–835, Online at http://onlinelibrary.wiley.com/doi/10.1002/2013GL058374/pdf.

Philpotts AR (1964) Origin of pseudotachylites. *American Journal of Science* 262: 1008–1035.

Platt JD and Rice JR (2013) Migration of the deforming zone during seismic shear and implications for field observations, dynamic weakening, and the onset of melting. In: *AGU Fall Meeting,* Abstracts T33C-2646.

Power WL and Tullis TE (1989) The relationship between slickenside surfaces in fine-grained quartz and the seismic cycle. *Journal of Structural Geology* 11: 879–893.

Power WL and Tullis TE (1992) The contact between opposing fault surfaces at Dixie Valley, Nevada, and implications for fault mechanics. *Journal of Geophysical Research* 97: 15425–15435.

Prakash V and Clifton RJ (1993) Time resolved dynamic friction measurements in pressure-shear. In: *Experimental Techniques in the Dynamics of Deformable Bodies*, pp. 33–47. New York: ASME.

Prakash V and Yuan F (2004) Results of a pilot study to investigate the feasibility of using new experimental techniques to measure sliding resistance at seismic slip rates. *EOS, Transactions, American Geophysical Union* 85(47), Fall Meeting Supplement, Abstract T21D-02.

Raj R and Ashby MF (1971) On grain boundary sliding and diffusional creep. *Metallurgical Transactions* 2: 1113–1127.

Rajagopalan S and Prakash V (1999) A modified Kolsky bar for investigating dynamic friction. *Experimental Mechanics* 39(4): 295–303.

Ranjith K and Rice J (2001) Slip dynamics at an interface between dissimilar materials. *Journal of the Mechanics and Physics of Solids* 49: 341–361.

Rapoport L, Leshchinsky V, Lvovsky M, et al. (2003) Superior tribological properties of powder materials with solid lubricant nanoparticles. *Wear* 255(7–12): 794–800.

Reches Z and Lockner DA (2010) Fault weakening and earthquake instability by powder lubrication. *Nature* 467: 452–455. http://dx.doi.org/10.1038/nature09348.

Rempel AW and Rice JR (2006) Thermal pressurization and onset of melting in fault zones. *Journal of Geophysical Research: Solid Earth* 111(B9): B09314. http://dx.doi.org/10.01029/02006JB004314.

Rice JR (1992) Fault stress states, pore pressure distributions and the weakness of the San Andreas Fault. In: Evans B and Wong T (eds.) *Fault Mechanics and Transport Properties of Rocks*, pp. 475–503. San Diego, CA: Academic Press.

Rice JR (1999) Flash heating at asperity contacts and rate-dependent friction. *EOS, Transactions American Geophysical Union* 80: F681.

Rice JR (2006) Heating and weakening of faults during earthquake slip. *Journal of Geophysical Research: Solid Earth* 111(B5). http://dx.doi.org/10.1029/2005JB004006.

Rice JR and Cocco M (2007) Seismic fault rheology and earthquake dynamics. In: Handy MR (ed.) *The Dynamics of Fault Zones*, pp. 99–137. Cambridge, MA: MIT Press.

Roig Silva C, Goldsby DL, Di Toro G, and Tullis TE (2004a) The role of silica content in dynamic fault weakening due to gel lubrication. In: *Paper presented at 2004 SCEC Annual Meeting Proceedings and Abstracts, Southern California Earthquake Center, Palm Springs, California.*

Roig Silva C, Goldsby DL, Di Toro G, and Tullis TE (2004b) The role of silica content in dynamic fault weakening due to gel lubrication. *EOS, Transactions, American Geophysical Union* 85(47), Fall Meeting Supplement, T21D-07.

Rosakis AJ (2002) Intersonic shear cracks and fault ruptures. *Advances in Physics* 51(4): 1189–1257.

Rudnicki JW and Rice JR (2006) Effective normal stress alteration due to pore pressure changes induced by dynamic slip propagation on a plane between dissimilar materials. *Journal of Geophysical Research: Solid Earth* 111: B10308, 18 p, http://dx.doi.org/10.1029/2006JB004396.

Ruina AL (1983) Slip instability and state variable friction laws. *Journal of Geophysical Research* 88: 10359–10370.

Sammis C, Lockner DA, and Reches Z (2011) The role of adsorbed water on the friction of a layer of submicron particles. *Pure and Applied Geophysics* 168: 2325–2334.

Schubnel A, Brunet F, Hilairet N, Gasc J, Wang Y, and Green HW II (2013) Deep-focus earthquake analogs recorded at high pressure and temperature in the laboratory. *Science* 341(6152): 1377–1380. http://dx.doi.org/10.1126/science.1240206.

Scruggs VJ and Tullis TE (1998) Correlation between velocity dependence of friction and strain localization in large displacement experiments on feldspar, muscovite and biotite gouge. *Tectonophysics* 295: 15–40.

Segall P and Rice JR (1995) Dilatancy, compaction, and slip instability of a fluid infiltrated fault. *Journal of Geophysical Research* 100: 22155–22173.

Sengers JV and Watson JTR (1986) Improved international formulations for the viscosity and thermal conductivity of water substance. *Journal of Physical and Chemical Reference Data* 15(4): 1291–1314.

Serdobintseva VV and Kalinin DV (2001) Mechanism of reversible thixotropic transformation in concentrated suspensions of monodisperse spherical silica particles: Analysis in terms of colloidal crystal formation and growth. *Colloidal Journal* 63(5): 626–629.

Serdobintseva VV, Kalinin DV, and Vosel SV (1998) The forms of colloid silica participating in the formation of noble opal and the mechanism of silicification of its gel crystals. *Geologiya I Geofizika* 39(8): 1116–1120.

Seront B, Wong TF, Caine JS, Forster CB, Bruhn RL, and Fredrich JT (1998) Laboratory characterization of hydromechanical properties of a seismogenic normal fault system. *Journal of Structural Geology* 20(7): 865–881.

Shaw HR (1972) Viscosities of magmatic silicate liquids: An empirical method of prediction. *American Journal of Science* 272: 870–889.

Shaw DJ (1980) *Introduction to Colloid and Surface Chemistry*, p. 273. London: Butterworths.

Shi Z and Ben-Zion Y (2006) Dynamic rupture on a bimaterial interface governed by slip-weakening friction. *Geophysical Journal International* 165: 469–484. http://dx.doi.org/10.1111/j.1365-1246X.2006.02853.x.

Shimamoto T and Lin A (1994) Is frictional melting equilibrium melting or non-equilibrium melting? (in Japanese with English abstract). *Journal of Structural Geology* 39: 79–84.

Shimamoto T and Tsutsumi A (1994) A new rotary-shear high-velocity frictional testing machine: Its basic design and scope of research (in Japanese with English abstract). *Journal of Structural Geology* 39: 65–78.

Sibson RH (1973) Interactions between temperature and fluid pressure during earthquake faulting – A mechanism for partial or total stress relief. *Nature* 243: 66–68.

Sibson RH (1975) Generation of pseudotachylyte by ancient seismic faulting. *Geophysical Journal of the Royal Astronomical Society* 43: 775–794.

Sibson RH (2003) Thickness of the seismic slip zone. *Bulletin of the Seismological Society of America* 93: 1169–1178.

Sibson RH and Toy VG (2006a) The habitat of fault-generated pseudotachylyte: Presence vs. absence of friction melt. In: Abercrombie R (ed.) *Earthquakes: Radiated Energy and the Physics of Faulting. Geophysical Monograph Series*, vol. 17, pp. 153–166. Washington, DC: American Geophysical Union. http://dx.doi.org/10.1029/1170GM1016.

Sibson RH and Toy V (2006b) The habitat of fault-generated pseudotachylyte: Presence vs. absence of friction melt. In: Abercrombie R, McGarr A, Kanamori H, and DiToro G (eds.) *Radiated Energy and the Physics of Earthquake Faulting. Geophysical Monograph Series*, vol. 170, pp. 43–57. Washington, DC: AGU.

Spray JG (1987) Artificial generation of pseudotachylite using friction welding apparatus: Simulation of melting on a fault plane. *Journal of Structural Geology* 9: 49–60.

Spray JG (1988) Generation and crystallization of an amphibolite shear melt – An investigation using radial friction welding apparatus. *Contributions to Mineralogy and Petrology* 99(4): 464–475.

Spray JG (1989a) Slickenside formation by surface melting during the mechanical excavation of rock. *Journal of Structural Geology* 11(7): 895–905.

Spray JG (1989b) Special issue – Friction phenomena in rock – An introduction. *Journal of Structural Geology* 11(7): 783–785.

Spray JG (1992) A physical basis for the frictional melting of some rock forming minerals. *Tectonophysics* 204: 205–221.

Spray JG (1993) Viscosity determinations of some frictionally generated silicate melts – Implications for fault zone rheology at high-strain rates. *Journal of Geophysical Research* 98: 8053–8068.

Spray JG (1995) Pseudotachylyte controversy – Fact or friction. *Geology* 23(12): 1119–1122.

Spray JG (2005) Evidence for melt lubrication during large earthquakes. *Geophysical Research Letters* 32: Art. No. L1007,1301. http://dx.doi.org/10.1029/2004GL1022,1293.

Spray JG and Thompson LM (1995) Friction melt distribution in a multiring impact basin. *Nature* 373(6510): 130–132.

Stel H (1981) Crystal-growth in cataclasites – Diagnostic microstructures and implications. *Tectonophysics* 78(1–4): 585–600.

Stel H and Lankreyer AC (1994) Flow and deformation of viscous, silica-oversaturated dispersions in low-grade faults. *Journal of Structural Geology* 16(3): 303–313.

Sulem J, Vardoulakis I, Ouffroukh H, Boulon M, and Hans J (2004) Experimental characterization of the thermo-poro-mechanical properties of the Aegion Fault gouge. *Comptes Rendus Geoscience* 336(4–5): 455–466.

Swanson MT (1988) Pseudotachylyte-bearing strike-slip duplex structures in the fort foster brittle zone, S Maine. *Journal of Structural Geology* 10(8): 813–828.

Swanson MT (1990) Extensional duplexing in the York cliffs strike-slip-fault system, southern coastal Maine. *Journal of Structural Geology* 12(4): 499–512.

Swanson MT (1992) Fault structure, wear mechanisms and rupture processes in pseudotachylyte generation. *Tectonophysics* 204: 223–242.

Takeo M, Ide S, and Yoshida Y (1993) The 1993 Kushiro-Oki, Japan, Earthquake – A high stress-drop event in a subducting slab. *Geophysical Research Letters* 20(23): 2607–2610.

Titone B, Sayre K, Di Toro G, Goldsby DL, and Tullis TE (2001) The role of water in the extraordinary frictional weakening of quartz rocks during rapid sustained slip. *EOS, Transactions American Geophysical Union* 82: T31B–T0841B.

Tsutsumi A and Shimamoto T (1994) An attempt to measure the temperature of frictional melts of rocks produced during rapid fault motion (in Japanese with English abstract). *Journal of Structural Geology* 39: 103–114.

Tsutsumi A and Shimamoto T (1996) Frictional properties of monzodiorite and gabbro during seismogenic fault motion. *Journal of Geological Society of Japan* 102: 240–248.

Tsutsumi A and Shimamoto T (1997a) High-velocity frictional properties of gabbro. *Geophysical Research Letters* 24: 699–702.

Tsutsumi A and Shimamoto T (1997b) Temperature measurements along simulated faults during seismogenic fault motion. In: Hong Y (ed.) *Contemporary Lithospheric Motion Seismic Geology*, Proceedings of the 30th International Geological Congress, vol. 5, pp. 223–232. Leiden: VSP.

Tullis TE (1988) Rock friction constitutive behavior from laboratory experiments and its implications for an earthquake prediction field monitoring program. *Pure and Applied Geophysics* 126: 555–588.

Tullis TE (1996) Rock friction and its implications for earthquake prediction examined via models of Parkfield earthquakes. In: Knopoff L (ed.) *Earthquake Prediction: The Scientific Challenge*, *Proceedings of the National Academy of Sciences of America,* pp. 3803–3810.

Tullis TE (2014) Low strength of major faults may control the orientation of stresses. *Nature,* in preparation.

Tullis TE and Goldsby DL (2003) Flash melting of crustal rocks at almost seismic slip rates. *EOS, Transactions, American Geophysical Union* 84(46), Fall Meeting Supplement, Abstract S51B-05.

Tullis TE and Goldsby DL (2013) Laboratory experiments on rock friction focused on understanding earthquake mechanics. *Final Technical Report for Grant Number: G12AP20080*, http://earthquake.usgs.gov/research/external/research.php.

Tullis TE and Weeks JD (1986) Constitutive behavior and stability of frictional sliding of granite. *Pure and Applied Geophysics* 124: 10–42.

Tullis TE, et al. (2007) Rheology of fault rocks and their surroundings. In: Handy M, Hirth G, and Hovius N (eds.) *Tectonic Faults – Agents of Change on a Dynamic Earth*, pp. 183–204. Cambridge, MA: MIT Press.

Ujiie K, et al. (2013) Low coseismic shear stress on the Tohoku-Oki megathrust determined from laboratory experiments. *Science* 342: 1211–1214.

van der Elst NJ, Brodsky EE, Le Bas P-Y, and Johnson PA (2012) Auto-acoustic compaction in steady shear flows: Experimental evidence for suppression of shear dilatancy by internal acoustic vibration. *Journal of Geophysical Research* 117: B09314. http://dx.doi.org/10.1029/2011JB008897.

Weertman J (1963) Dislocations moving uniformly on the interface between isotropic media of different elastic properties. *Journal of the Mechanics and Physics of Solids* 11: 197–204.

Weertman J (1980) Unstable slippage across a fault that separates elastic media of different elastic constants. *Journal of Geophysical Research* 85: 1455–1461.

Weertman J (2002) Subsonic type earthquake dislocation moving at approximately square root 2x shear wave velocity on interface between half spaces of slightly different elastic constants. *Geophysical Research Letters* 29; 109-1–109-4. http://dx.doi.org/10.1029/2001GL013916.

Wenk HR (1978) Are pseudotachylites products of fracture or fusion? *Geology* 6(8): 507–511.

Wenk HR (1979) Are pseudotachylites products of fracture or fusion – Reply. *Geology* 7(4): 162–163.

Wenk HR, Johnson LR, and Ratschbacher L (2000) Pseudotachylites in the eastern Peninsular Ranges of California. *Tectonophysics* 321(2): 253–277.

Wibberley CAJ (2002) Hydraulic diffusivity of fault gouge zones and implications for thermal pressurization during seismic slip. *Earth, Planets and Space* 54: 1153–1171.

Wibberley CAJ and Shimamoto T (2003) Internal structure and permeability of major strike slip fault zones: The Median Tectonic Line in Mid Prefecture, Southwest Japan. *Journal of Structural Geology* 25: 59–78.

Wibberley CAJ and Shimamoto T (2005) Earthquake slip weakening and asperities explained by thermal pressurization. *Nature* 436: 689–692.

Wornyoh EYA, Jasti VK, and Higgs CF (2007) A review of dry particulate lubrication: Powder and granular materials. *Journal of Tribology* 129(2): 438–449. http://dx.doi.org/10.1115/1111.2647859.

Xia K, Rosakis AJ, Kanamori H, and Rice JR (2005) Laboratory earthquakes along inhomogeneous faults: Directionality and supershear. *Science* 308: 681–684.

Yund RA, Blanpied ML, Tullis TE, and Weeks JD (1990) Amorphous material in high strain experimental fault gouges. *Journal of Geophysical Research* 95: 15589–15602.

Zhang S and Tullis TE (1998) The effect of fault slip on permeability and its anisotropy in an artificial quartz gouge. *Tectonophysics* 295: 41–52.

Zhang S, Tullis T, and Scruggs V (1999) Permeability anisotropy and pressure dependency of permeability in experimentally sheared gouge materials. *Journal of Structural Geology* 21: 795–806.

Zhang S, Tullis TE, and Scruggs VJ (2001) Implications of permeability and its anisotropy in a mica gouge for pore pressures in fault zones. *Tectonophysics* 335: 37–50.

Zoback MD, et al. (1987) New evidence on the state of stress of the San Andreas Fault system. *Science* 238: 1105–1111.

This page intentionally left blank

4.07 The Role of Fault-Zone Drilling

MD Zoback, Stanford University, Stanford, CA, USA
S Hickman and W Ellsworth, US Geological Survey, Menlo Park, CA, USA

Published by Elsevier B.V.

This chapter is reproduced from the previous edition, Volume 4, pp. 649–674, © 2007 Elsevier B.V.

4.07.1	**Introduction: Why Drill to Study Earthquakes?**	161
4.07.2	**Fluids and Faulting**	162
4.07.2.1	Sources of Fault-Zone Fluids	162
4.07.2.2	Fault-Zone Permeability	163
4.07.2.3	Transient Fluid Pressure Effects	163
4.07.2.4	Chemical Effects of Fluids on Fault-Zone Rheology	163
4.07.3	**Frictional Strength of Faults**	164
4.07.3.1	Weak Plate-Bounding Faults	165
4.07.4	**Near-Field Observations of Earthquake Nucleation and Propagation**	167
4.07.5	**Fault-Zone Drilling Projects**	168
4.07.5.1	Fault Behavior	168
4.07.5.2	Fluid Pressure	168
4.07.5.3	Fault Fluids	169
4.07.5.4	Fault-Zone Properties and Physical Parameters	169
4.07.5.5	Fault Structure and Materials	169
4.07.5.6	San Andreas Fault Observatory at Depth	169
4.07.5.7	Nojima Fault	170
4.07.5.8	Chelungpu Fault	171
4.07.5.9	Gulf of Corinth	172
4.07.5.10	NELSAM	175
4.07.5.11	NanTroSeize	176
4.07.6	**Summary**	176
References		178

4.07.1 Introduction: Why Drill to Study Earthquakes?

The objective of fault-zone drilling projects is to directly study the physical and chemical processes that control deformation and earthquake generation within active fault zones. An enormous amount of field, laboratory, and theoretical work has been directed toward the mechanical and hydrologic behavior of faults over the past several decades. Nonetheless, it is currently impossible to differentiate between – or even adequately constrain – the numerous conceptual models of active faults proposed over the years. For this reason, the Earth science community is left in the untenable position of having no generally accepted paradigm for the mechanical behavior of faults at depth. One of the primary causes for this dilemma is the difficulty of either directly observing or inferring physical properties and deformation mechanisms along faults at depth, as well as the need to observe directly key parameters such as the state of stress acting on faults at depth, pore fluid pressure (and its possible variation in space and time), and processes associated with earthquake nucleation and rupture. Today, we know very little about the composition of active faults at depth, their constitutive properties, the state of in situ stress or pore pressure within fault zones, the origin of fault-zone pore fluids, or the nature and significance of time-dependent fault-zone processes.

Most of what we now know about the structure, composition, and deformation mechanisms of crustal faults has been learned from geologic investigations of exhumed faults, particularly where erosion has exposed previously deeply buried fault rocks. These field observations have proved to be particularly useful for several reasons. First, field observations of exhumed faults allow broad coverage with respect to variations in faulting style (e.g., comparing strike-slip, normal, and reverse faults), fault movement history, and local geology. Second, where sufficient surface outcrops exist, field observations can readily address issues related to geometric complexity and spatial heterogeneity in physical properties and fluid composition (e.g., Evans and Chester, 1995; Kerrich et al., 1984; Parry, 1994).

Constraints on the mechanical state and physical properties of active fault zones (e.g., fluid pressure, stress, and permeability) from surface observations are indirect and subject to alternate interpretations. Stress heterogeneities induced by fault slip can lead to considerable uncertainties in inferring past fluid pressures from observations of vein geometry in outcrop. In all of these investigations, a complex history of uplift and denudation may have severely altered evidence of deformation mechanisms, fault-zone mineralogy, and fluid composition operative during fault slip. This problem is especially acute for solution transport deformation mechanisms (e.g., pressure solution and crack healing/sealing) and other low-activation-energy processes, as the deformation microstructures formed at depth are easily overprinted by ongoing deformation as the fault rocks are brought to the surface. Thus, the importance of

fluids in earthquake generation and rupture is impossible to assess with any degree of certainty based solely on studies of exhumed fault rocks.

Fault-zone drilling projects allow us to address a number of first-order questions related to fault mechanics:

What are the mineralogy, deformation mechanisms, and constitutive properties of fault gouge? Why do some faults creep? What are the strength and frictional properties of recovered fault rocks at realistic in situ conditions of stress, fluid pressure, temperature, strain rate, and pore fluid chemistry? What determines the depth of the shallow seismic-to-aseismic transition? What do mineralogical, geochemical, and microstructural analyses reveal about the nature and extent of water–rock interaction?

What is the fluid pressure and permeability within and adjacent to fault zones? Are there superhydrostatic fluid pressures within some fault zones and through what mechanisms are these pressures generated and/or maintained? How does fluid pressure vary during deformation and episodic fault slip (creep and earthquakes)? Do fluid pressure seals exist within or adjacent to fault zones and at what scales?

What are the composition and origin of fault-zone fluids and gases? Are these fluids of meteoric, metamorphic, or mantle origin (or combinations of the three)? Is fluid chemistry relatively homogeneous, indicating pervasive fluid flow and mixing, or heterogeneous, indicating channelized flow and/or fluid compartmentalization?

How do stress orientations and magnitudes vary across fault zones? Are principal stress directions and magnitudes different within the deforming core of weak fault zones than in the adjacent (stronger) country rock, as predicted by some theoretical models? How does fault strength measured in the near field compare with depth-averaged strengths inferred from heat-flow and regional stress directions? What is the nature and origin of stress heterogeneity near active faults?

How do earthquakes nucleate? Does seismic slip begin suddenly or do earthquakes begin slowly with accelerating fault slip? Do the size and duration of this precursory slip episode, if it occurs, scale with the magnitude of the eventual earthquake? Are there other precursors to an impending earthquake, such as changes in pore pressure, fluid flow, crustal strain, or electromagnetic field?

How do earthquake ruptures propagate? Do earthquake ruptures propagate as a uniformly expanding crack or as a 'slip pulse'? What is the effective (dynamic) stress during seismic faulting? How important are processes such as shear heating, transient increases in fluid pressure, and fault-normal opening modes in lowering the dynamic frictional resistance to rupture propagation?

How do earthquake source parameters scale with magnitude and depth? What is the minimum size earthquake that occurs on faults? How is long-term energy release rate partitioned between creep dissipation, seismic radiation, dynamic frictional resistance, and grain size reduction (i.e., by integrating fault-zone monitoring with laboratory observations on core)?

What are the physical properties of fault-zone materials and country rock (seismic velocities, electrical resistivity, density, and porosity)? How do physical properties from core samples and downhole measurements compare with properties inferred from surface geophysical observations? What are the dilational, thermoelastic, and fluid transport properties of fault and country rocks, and how might they interact to promote either slip stabilization or transient overpressurization during faulting?

What processes control the localization of slip and strain? Are fault surfaces defined by background microearthquakes and creep the same? Would active slip surfaces be recognizable through core analysis and downhole measurements in the absence of seismicity and/or creep?

4.07.2 Fluids and Faulting

Among the many compelling reasons for drilling into active faults is the possibility to study the role of fluids, fluid pressure, and fluid flow in earthquake processes. A long-standing (and still-growing) body of evidence suggests that fluids are intimately linked to a variety of faulting processes (see review by Hickman et al., 1995). These include the long-term structural and compositional evolution of fault zones, fault creep, and the nucleation, propagation, arrest, and recurrence of earthquake ruptures.

The concept that high fluid pressures and the localization of deformation are often linked is widely accepted in the structural geology literature (e.g., Fyfe et al., 1978; Hubbert and Rubey, 1959) and has been reinforced by studies of active accretionary prisms in subduction complexes and their fossil equivalents (e.g., Dahlen, 1990; Fisher, 1996; Saffer and Bekins, 2006). Better understanding of the role of fluids in faulting in accretionary prisms is one goal behind the plan to drill in the Nankai subduction zone (the NanTroSeize project), as discussed in the succeeding text. A point to note is that there is good evidence that seismic rupturing, in at least some instances (e.g., the Western Taiwan fold and thrust belt and the western margin of the Great Valley adjacent to the San Andreas Fault), is occurring in fluid-overpressured crust (Davis et al., 1983; Sibson, 1990). The Taiwan Chelungpu Fault Drilling Project (TCDP), discussed later, is related to drilling into the fault responsible for the M 7.7 Chi-Chi earthquake of 1999, which occurred along the western margins of the Western Taiwan fold and thrust belt.

4.07.2.1 Sources of Fault-Zone Fluids

Potential sources of fluids in brittle faults and shear zones include metamorphic fluid generated by dehydration of minerals during prograde metamorphism (including shear heating), fluid trapped in pore space as sedimentary formation brines, meteoric water carried downward by circulation, and release of volatiles from molten magma or the upper mantle (e.g., Hacker et al., 1995; Kerrich et al., 1984; Ko et al., 1997; Wakita and Sano, 1987). The high fluid pressures that have been postulated within the San Andreas Fault zone (being tested as part of the San Andreas Fault Observatory at Depth (SAFOD) drilling project discussed later) might be generated and maintained by continued upwelling of overpressured fluids within the fault zone and leakage of these fluids into

the country rock (Rice, 1992). Alternatively, high fluid pressures might result from the sealing of locally derived high-pressure fluids within the fault zone once pressure gradients drop below a critical 'threshold' required to overcome forces between molecular water and mineral surfaces in very small cracks and pores (Byerlee, 1990).

Kennedy et al. (1997) argued that the elevated $^{3}He/^{4}He$ ratios they observed in springs and wells located along a broad zone encompassing the San Andreas Fault system indicate that significant quantities of mantle-derived fluids are entering the fault zone through the ductile lower crust at near-lithostatic pressure. However, without direct sampling of fluids from within the San Andreas Fault zone at depth, it is unclear whether these fluids are ascending through a broad, fractured, and faulted zone associated with the overall plate boundary or are narrowly focused within the (permeable) core of the San Andreas Fault itself and hence intimately involved in the physics of faulting as envisioned by Rice (1992) and others.

4.07.2.2 Fault-Zone Permeability

The permeability structure of shear zones and brittle faults has recently been the focus of field studies that both confirm and extend observations made years ago by mining geologists. Large faults are not discrete surfaces, but rather are a braided array of slip surfaces encased in a highly fractured and often hydrothermally altered transition or 'damage' zone (Bruhn et al., 1990, 1994; Chester et al., 1993; Schulz and Evans, 2000; Smith et al., 1990). Structural and mineralogical textures indicate that episodic fracturing and brecciation are followed by cementation and crack healing, leading to cycles of permeability enhancement and reduction accompanied by episodic fluid flow along faults (e.g., Eichhubl and Boles, 2000).

Theoretical modeling (Sleep, 1995; Sleep and Blanpied, 1992, 1994) showed that the generation of dilatant pores and microcracks during earthquakes in a hydraulically isolated fault zone, followed by creep compaction between earthquakes, might lead to cyclically high fluid pressures along faults. Miller (1996) and Fitzenz and Miller (2003) had used the Sleep and Blanpied (1992) model to numerically simulate temporal variations in fluid pressure within faults and their role in controlling earthquake interactions and periodicity, making assumptions about the rheological and hydrologic properties of fault and country rocks. However, using these models to predict the behavior of real faults in the Earth will require direct in situ measurements and sampling in active fault zones at seismogenic depths.

A possibly relevant development from studies of fluid pressure in sedimentary basins has been the revelation from borehole measurements of abrupt transitions, both vertically and laterally, between distinct fluid pressure regimes in some sedimentary basins. These 'fluid pressure compartments' are bounded by seals, which in some cases are stratigraphic (e.g., shale horizons) but in others are gouge-rich faults or thin zones of hydrothermal cementation, which cut across stratigraphy (Dewers and Ortoleva, 1994; Hunt, 1990; Martinsen, 1997; Powley, 1990). By analogy with these observations, Byerlee (1993) proposed a model in which contiguous vertical and horizontal seals within a fault zone would lead to discrete fluid pressure compartments (i.e., tabular lenses), the rupture of which might be important in earthquake nucleation and propagation (see Lockner and Byerlee, 1995). Although direct evidence for these fault-zone fluid compartments in active fault zones is lacking, negative polarity reflections (bright spots) on seismic reflection images acquired over some accretionary prisms have been interpreted to indicate the existence of high-pressure fluid compartments along the basal décollements (Bangs et al., 2004; Moore and Vrolijk, 1992; Moore et al., 1995; Shipley et al., 1994).

4.07.2.3 Transient Fluid Pressure Effects

A range of physical effects arising from the mechanical response of fluid-saturated crust has been invoked to account for time-dependent phenomena associated with faulting such as slow earthquakes, creep events, earthquake swarms, and aftershock activity and its decay (e.g., Dreger et al., 2000; Hainzl and Ogata, 2005; Nur and Booker, 1972; Rice and Cleary, 1976). Transient changes in fluid pressure and effective stress have also been suggested to play a direct role in rupture propagation and arrest. Shear resistance on the rupture surface may be dramatically lowered by localized increases in fluid pressure from frictional heating or locally elevated as a consequence of pore fluid diffusion and dilatant hardening at fault jogs and other irregularities (Andrews, 2002; Lachenbruch, 1980; Mase and Smith, 1987; Rudnicki, 1988; Segall and Rice, 1995; Sibson, 1973, 1985; Sleep, 1995). Continuous monitoring of fluid pressure within active, seismogenic faults is a critical objective of both the SAFOD and Gulf of Corinth scientific drilling projects discussed in the succeeding text.

4.07.2.4 Chemical Effects of Fluids on Fault-Zone Rheology

Over the past several years, a number of fault mechanics models have been either developed or refined that incorporate solution transport deformation mechanisms that may weaken and/or destabilize the fault zone. However, complicating this issue enormously is the fact that under only slightly varied environmental and mineralogical conditions, similar processes can act to cement the fault zone together, thereby increasing fault strength (see Hickman and Evans, 1992). The experimental and theoretical studies on which these models are based are now focusing on processes that have long been inferred as being important from field observations of natural fault and shear zones, such as pressure solution, fluid-assisted retrograde mineral reactions, crack healing, and cementation (e.g., Boullier and Robert, 1992; Bruhn et al., 1990; Chester et al., 1993; Eichhubl and Boles, 2000; Kerrich et al., 1984; Power and Tullis, 1989). These deformation mechanisms are all interrelated, in that they depend upon thermally activated chemical reactions between the rock and pore fluid as well as the rates at which dissolved species are transported through the pore fluid.

Laboratory and theoretical investigations have shown that pressure solution may be important in reducing long-term fault strength and in promoting aseismic slip (i.e., creep) along faults (e.g., Bos and Spiers, 2001; Chester and Higgs, 1992; Rutter and Mainprice, 1979). In contrast, solution transport processes such as crack healing and sealing and cementation may cause the

welding together of asperities or fault gouge, leading to time-dependent fault strengthening between earthquakes (e.g., Angevine et al., 1982; Hickman and Evans, 1992; Karner et al., 1997). Laboratory friction experiments conducted under hydrothermal conditions suggest that a change in dominant deformation mechanism with increasing depth from brittle deformation to solution transport creep might control the depth at which the seismic-to-aseismic transition occurs in the crust (Blanpied et al., 1991). Observations of aligned fibrous serpentine in an exhumed branch of the San Andreas Fault system suggest that dissolution–diffusion–crystallization processes may be important in reducing fault strength and promoting aseismic slip (creep) along faults (Andreani et al., 2005).

Hydrothermal mineral reactions can also weaken crustal rocks when the reaction products are weaker than the reactants (see Wintsch et al., 1995). Based upon observations of exhumed shear zones in granite, Janecke and Evans (1988) argued that muscovite formed from the breakdown of feldspar might dramatically lower the ductile shear strength of the granite, even at temperatures well below those necessary for the plastic flow of quartz. At least at shallow depths, fault zones such as the San Andreas are mostly composed of clay- and mica-rich gouge resulting from the hydrolysis of feldspar (e.g., Wu, 1978), suggesting an enhancement of the feldspar breakdown reaction within the fault zone. Reactions in the olivine–talc–serpentine–water system have been demonstrated to dramatically lower the shear strength of ultramafic rocks in laboratory friction experiments (Pinkston et al., 1987).

4.07.3 Frictional Strength of Faults

In addition to studying the role of fluids in faulting, fault-zone drilling provides the opportunity to directly measure the state of stress acting within and adjacent to active faults. As reviewed by Zoback and Healy (1984), Hickman (1991), and Townend and Zoback (2000), in situ stress measurements in a variety of faulting regimes, in conjunction with information on the attitude of nearby active faults, indicate fault strengths in intraplate areas that are comparable to those predicted by combining Coulomb faulting theory and laboratory-derived coefficients of friction between 0.6 and 1.0 (Byerlee, 1978). This is sometimes known as Byerlee's law (Brace and Kohlstedt, 1980). When pore pressure is near hydrostatic (as has found to be the case at all sites where deep drilling has occurred in crystalline rock (see review in Townend and Zoback (2000)), high stress magnitudes, sometimes referred to as 'hydrostatic' Byerlee's law, are predicted. The high stress at depth consistent with Byerlee's law has been measured at a number of sites, including the Rocky Mountain Arsenal, Denver (Healy et al., 1968); Rangely, Colorado (Raleigh et al., 1972; Zoback and Healy, 1984); the Nevada Test Site (Stock et al., 1985); the Fenton Hill geothermal site, New Mexico (Barton et al., 1988; Fehler, 1989); Moodus, Connecticut (Baumgärtner and Zoback, 1989; Mrotek et al., 1988); Dixie Valley, Nevada (Hickman et al., 1997); the Siljan deep borehole in Sweden (Lund and Zoback, 1999); and, to ~8 km depth in the KTB drilling project, Oberpfalz, West Germany (Brudy et al., 1997). Thus, Byerlee's law (Brace and Kohlstedt, 1980), which was established on the basis of simple faulting theory and laboratory friction experiments (Byerlee, 1978), appears valid for faults within plate interiors. Studies of lithospheric flexure in response to sediment, volcanic, and internal loads (e.g., Kirby, 1983; McNutt, 1980; McNutt and Menard, 1982) also indicate that differential stresses in much of the Earth's crust are high and can approach the magnitude of stress predicted by Byerlee's law. Hence, such sources of stress are capable of generating crustal stresses large enough to induce faulting.

Figure 1 (from Townend and Zoback (2000)) shows a compilation of stress measurements in relatively deep wells and boreholes in various parts of the world. As shown, the ratio of the maximum and minimum effective stresses corresponds to a crust in frictional failure equilibrium with a coefficient of friction ranging between 0.6 and 1.0. The rate at which stress increases with depth is based on the assumption that the ratio of maximum to minimum effective stress is

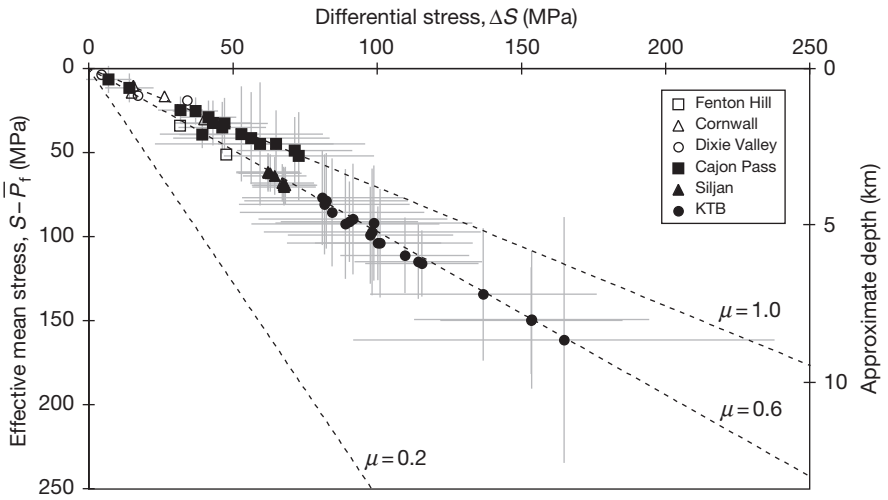


Figure 1 Stress measurements in deep boreholes indicate high crustal strength in accordance with Coulomb faulting theory with laboratory-determined coefficients of friction between 0.6 and 1.0 and hydrostatic pore pressure, which is observed in each borehole. It should be noted that the uncertainty estimates in the figure are likely significantly smaller than those shown. Reproduced from Townend J and Zoback MD (2000) How faulting keeps the crust strong. *Geology* 28(5): 399–402.

limited by ratio of shear to normal stress on preexisting faults well oriented to slip in the current stress field (see Zoback and Healy, 1984). Pore pressure was observed to be essentially equal to hydrostatic pressure at all the sites where deep drilling into crystalline rocks in the crust has taken place. Because stress magnitudes increase with depth due to increasing overburden stress, mean effective stress (mean stress minus the pore pressure, P_f) at depth scales approximately with depth as shown in the figure and discussed by Townend and Zoback (2000).

There are two implications of the data shown in **Figure 1**. First, Byerlee's law, defined on the basis of hundreds of laboratory experiments, appears to be applicable to faults in situ. The consistency of in situ friction coefficients and those measured in the lab is a rather amazing result when one considers the huge difference between the size of samples used for friction experiments in the lab and the size of real faults in situ, the variability of roughness of the sliding surface, and variability in the rock types encountered in the field and studied in the lab. Second, everywhere that stress magnitudes have been measured at appreciable depth, they indicate that they are controlled by the frictional strength of preexisting faults in the crust. In other words, the Earth's crust appears to be in a state of frictional failure equilibrium and the law that describes that state is simple Coulomb friction.

4.07.3.1 Weak Plate-Bounding Faults

A substantial (and growing) body of evidence indicates that slip in crustal earthquakes along major plate-bounding faults (like the San Andreas) occurs at much low levels of shear stress than those shown in **Figure 1**. This hypothesis (sometimes referred to as the San Andreas stress/heat flow paradox, e.g., Lachenbruch and Sass (1980, 1992) and Zoback et al. (1987); see review by Hickman (1991)). In the context of the San Andreas, there are two principal lines of evidence that indicate that the fault has low frictional strength – the absence of frictionally generated heat and the orientation of the maximum principal stress in the crust adjacent to the fault. A large number of heat-flow measurements show no evidence of frictionally generated heat adjacent to the San Andreas Fault (Brune, 1970; Brune et al., 1969; Henyey and Wasserburg, 1971; Lachenbruch and Sass, 1980, 1992; Williams et al., 2004), which implies that shear motion along the fault is resisted by shear stresses approximately a factor of 5 less than those shown in **Figure 1**. Saffer et al. (2003) showed that it is highly unlikely that topographically driven fluid flow has an appreciable effect on these heat-flow measurements, indicating that the lack of frictionally generated heat in the vicinity of the San Andreas Fault is indeed indicative of average shear stress levels acting on the fault at depth.

In addition to the heat-flow data, the orientation of principal stresses in the vicinity of the fault also indicates that right-lateral strike-slip motion on the fault occurs in response to low levels of shear stress (Mount and Suppe, 1987; Oppenheimer et al., 1988; Townend and Zoback, 2001, 2004; Zoback and Beroza, 1993; Zoback et al., 1987). **Figure 2** (from Townend and Zoback, 2004) shows the direction of relative motion of the Pacific Plate with respect to the North American Plate

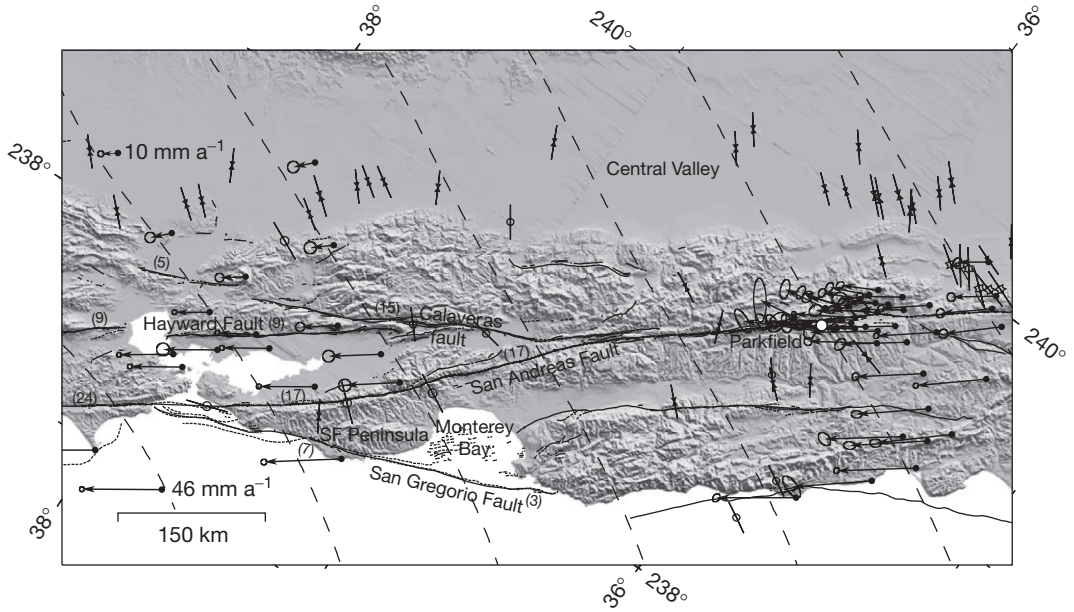


Figure 2 Maximum horizontal compressive stress (S_{Hmax}) and crustal velocity data from Central California shown in an oblique Mercator projection about the North America–Pacific Euler pole. S_{Hmax} directions determined from borehole breakouts are shown by inward-pointing arrows, from hydraulic fracturing experiments by short line segments with stars and from earthquake focal mechanism inversions by short line segments with split circles. The dashed trajectories show regional S_{Hmax} directions calculated using a model of lithospheric buoyancy and plate interaction by Flesch et al. (2000). The vectors illustrate crustal velocity data relative to North America. The numbers in parentheses are Quaternary fault slip rates. The location of the SAFOD drillhole is shown by the open circle near Parkfield. Reproduced from Murray MH and Segall P (2001) Modeling broadscale deformation in northern California and Nevada from plate motions and elastic strain accumulation. *Geophysical Research Letters* 28(22): 4315–4318; Townend J and Zoback MD (2004) Regional tectonic stress near the San Andreas Fault in Central and Northern California. *Geophysical Research Letters* 31(13): L15S11.

indicated by GPS measurements with blue arrows (the map projection is about the pole of relative motion) and the direction of maximum horizontal stress (black, inward pointed arrows). The direction of maximum horizontal compression is clearly at a very high angle to the fault, resulting in low levels of resolved shear stress on the fault.

The significance of regional stress orientation for the state of stress on the fault at depth has been challenged by Scholz (2000) who argued that if stress orientations near the fault rotate to being ~45° from the strike of the fault, both the fault and the crust could have comparably high frictional strength. As summarized by Zoback (2000), there is appreciable evidence that even close to the fault, the direction of maximum horizontal compression is at a high angle to the fault (Oppenheimer et al., 1988; Townend and Zoback, 2001, 2004; Zoback and Beroza, 1993). In fact, stress measurements in the SAFOD pilot hole (Hickman and Zoback, 2004) located only 1.8 km from the San Andreas Fault in Central California (discussed in the succeeding text) provide still more evidence in support of the hypothesis that the San Andreas is a weak fault imbedded in a strong crust. In other words, stress magnitudes in the crust adjacent to the fault are high but shear stresses resolved on the fault are low.

Horizontal differential stress magnitudes in the pilot hole were found to be high adjacent to the fault (i.e., consistent with measurements shown in **Figure 1**), while the orientation of maximum horizontal stress at depth in the hole indicates low levels of shear stress resolved onto the San Andreas Fault itself (Hickman and Zoback, 2004). Shear-wave anisotropy measurements in the SAFOD main hole further indicate that the direction of maximum horizontal compression remains nearly perpendicular to the fault to within a few hundred meters of the active fault trace at depth (Boness and Zoback, 2006). In addition, heat-flow measurements (to 2 km depth) in the SAFOD pilot hole are consistent with shallower data in the region, which show no evidence of frictionally generated heat (Williams et al., 2004).

There is mounting evidence that a number of other plate-boundary faults are similarly weak (see Ben-Avraham and Zoback, 1992; Kanamori, 1980; Lachenbruch and Thompson, 1972; Magee and Zoback, 1993; Mount and Suppe, 1992; Wang et al., 1995; Wilcock et al., 1990, and review by Hickman (1991)). The concept that plate-bounding faults like the San Andreas are weak faults imbedded in a strong crust has become widely accepted in recent years, and earthquake researchers are now faced with the problem of explaining why major plate-boundary faults are substantially weaker than the surrounding crust. In fact, the question of how crustal faults lose their strength is critically important in crustal mechanics and earthquake hazard reduction for a number of reasons:

- As the weakness of plate boundaries (relative to plate interiors) is a fundamental aspect of plate tectonics, how and why plate-boundary faults lose their strength is of fundamental global importance for understanding where plate boundaries form, how they evolve with time, and how deformation is partitioned along them.
- The roles of fluid pressure, intrinsic rock friction, chemical reactions, and the physical state of active fault zones in controlling fault strength must be known in order to simulate earthquakes in the laboratory and on the computer using representative fault-zone properties and physical conditions. This information will also allow for improved models of static stress transfer and earthquake triggering at a regional scale and between specific faults, as needed for intermediate-term seismic hazard forecasting following large earthquakes.
- Through long-term fault-zone monitoring and in situ observations of the earthquake source, it will be possible to improve models for earthquake rupture dynamics, including such effects as transient changes in fluid pressure, fault-normal opening modes, and variations in slip pulse duration. These observations can be used directly in attempts to generate improved predictions of near-field strong ground motion (amplitude, frequency content, and temporal characteristics) and more reliable models for dynamic stress transfer and rupture propagation. The latter processes are believed to control earthquake size (i.e., whether or not a small earthquake will grow into a large one) and, hence, are crucial to long-term probabilistic assessments of earthquake hazard.
- The results of fault-zone drilling experiments are critical to the development of more realistic models for the seismic cycle and assessment of the practicality of short-term earthquake prediction in two ways. During fault-zone monitoring, it will be possible to determine if earthquakes are preceded by accelerating fault slip (e.g., a nucleation phase) and/or transient changes in fluid pressure. Second, we will be able to determine whether or not factors that might dramatically lower fault strength (e.g., high pore pressure and/or chemical fluid–rock interactions) are closely related to the processes controlling earthquake nucleation. Our current knowledge of fault-zone processes is so poor that not only are we unable to make reliable short-term earthquake predictions but also we cannot scientifically assess whether or not such predictions are even possible.

In summary, while essentially all available data indicate that the frictional strength of intraplate crust is high, the frictional strength of plate-bounding faults is anomalously low. Taken together, the heat-flow data and the directional constraint (i.e., S_{Hmax} at 65–85° to the San Andreas Fault) suggest that the San Andreas Fault is weak in both an absolute and relative sense. In other words, the average shear stress required to cause faulting is low (comparable to earthquake stress drops), whereas the frictional strength of faults in the adjacent crust is much higher, consistent with Byerlee's law. Despite the fundamental nature of this finding, we have no direct in situ evidence indicating why this might be so, whether the mechanisms responsible for low strength along the San Andreas are likely to be found in other major fault systems or what role that these mechanisms might play in the processes of earthquake nucleation and propagation.

Numerous theories have been proposed over the past decade that are related to the weakness of the San Andreas and other plate-bounding faults. Knowledge of the in situ frictional properties of the San Andreas and other major active faults is not only of considerable scientific interest but also

critical for assessing the nature and potential magnitude of static stress transfer and earthquake triggering following large-to-intermediate size earthquakes.

Although the causes for the weakness of these faults are unknown, four general classes of explanations have been suggested:

- Elevated fluid pressure (as discussed earlier)
- Intrinsically low coefficients of friction of fault-zone materials
- Solution transport reactions and related low-stress deformation mechanisms (also discussed earlier)
- Dynamic weakening mechanisms such as shear heating/thermal pressurization or acoustic fluidization

If the coefficient of friction, μ, is equal to 0.6–0.9 on the San Andreas Fault, as predicted by Byerlee's law, then the heat-flow constraint could be satisfied if the in situ pore pressure, P_p, is greater than twice hydrostatic (Lachenbruch and Sass, 1980, 1992). However, if one assumes that principal stress magnitudes are constant across the fault zone and that $\mu \geq 0.6$, then high fluid pressures alone cannot explain the directional constraint as P_p would exceed the least principal stress once the angle between S_{Hmax} and the fault exceeds about 60° (e.g., Lachenbruch and McGarr, 1990; Scholz, 1989; Zoback et al., 1987). It has been suggested that large-scale yielding could lead to an increase in the magnitudes of the principal stresses within the fault zone relative to their values immediately outside of the fault (Rice, 1992). If so, this would allow P_p within the fault zone to exceed significantly the external magnitude of the least principal stress (Byerlee, 1990; Rice, 1992). In this manner, permanently high pore pressures within an intrinsically strong (i.e., high coefficient of friction) San Andreas Fault zone, in conjunction with much lower fluid pressures in the surrounding rock, could lower the fault strength sufficiently to satisfy both the heat-flow and directional constraints.

A model that is, in many respects, similar to that of Rice (1992) was proposed earlier by Byrne and Fisher (1990) to explain the apparent weakness of the basal décollement beneath the Kodiak accretionary prism in Southwest Alaska. Magee and Zoback (1993) applied the Rice model to explain the low frictional strength of the subduction zone associated with the $M \sim 8.2$ Tokachi-oki earthquake off northern Honshu, Japan. Although not requiring localized increases in stress magnitude, Fournier (1996) had presented a model in which near-lithostatic fluid pressures may be maintained within the San Andreas Fault zone at depths greater than about 6–10 km if the maximum differential stress at these depths is quite low (~10–30 MPa) and the tensile strength of the rock outside the fault zone remains high due to pervasive crack healing at elevated temperatures.

Alternatively, if one assumes that the fault is optimally oriented with respect to the principal stresses and that fluid pressures are hydrostatic, then the heat-flow constraint can be satisfied if the friction coefficient along the fault is less than about 0.2 (Lachenbruch and Sass, 1992). Similarly, at least in Central California where S_{Hmax} is at about 75–85° to the San Andreas Fault, the heat-flow and directional constraints can be simultaneously satisfied under conditions of uniformly hydrostatic fluid pressures if the friction coefficient is extremely low – about 0.1 or less – along the fault, and Byerlee's law is applicable outside the fault zone (Lachenbruch and McGarr, 1990; Lachenbruch and Sass, 1992). It is often proposed that the presence of clays or other weak minerals along the San Andreas and other faults might lead to anomalously low frictional resistance (e.g., Janecke and Evans, 1988; Wintsch et al., 1995; Wu, 1978). This inference has been supported by laboratory sliding experiments on synthetic clay-rich fault gouges (e.g., Bird, 1984; Logan and Rauenzahn, 1987; Shimamoto and Logan, 1981; Wang et al., 1980), on synthetic serpentinite gouges (Reinen and Tullis, 1995; Reinen et al., 1994), and on synthetic laumontite gouge (Hacker et al., 1995). However, these experiments are all at low-to-moderate confining pressures and temperatures. In contrast, experiments on natural clay-rich fault gouges collected from the San Andreas at depths of <0.4 km (Morrow et al., 1982), on synthetic clay-rich fault gouges (Morrow et al., 1992), and on synthetic serpentinite gouges (Moore et al., 1997) at high temperatures and/or confining pressures and hydrostatic fluid pressures indicate coefficients of friction at in situ conditions that are too high to be reconciled with either the heat-flow or directional constraints. In addition, both natural and synthetic fault gouges deformed in the laboratory generally fail to exhibit the slip-weakening or velocity-weakening behavior required for the generation of earthquakes (e.g., Byerlee and Summers, 1976; Logan and Rauenzahn, 1987; Marone et al., 1990; Morrow et al., 1992; Reinen et al., 1994). Thus, the importance of these materials in the rheology of the San Andreas Fault at seismogenic depths is unclear.

Another important class of models that might explain the low long-term strength of seismically active segments of the San Andreas and other major faults has called upon processes directly associated with earthquake rupture propagation. These dynamic weakening mechanisms include shear heating during slip, leading to transient high fluid pressures (Andrews, 2002; Lachenbruch, 1980) or melting (Rice, 2006; Sibson, 1973; Spray, 1987); reductions in normal stress accompanying the propagation of dilational waves along the fault (Andrews and Ben-Zion, 1997; Brune et al., 1993); the fluidization of fault-zone materials due to the channeling of coseismic acoustic energy (Melosh, 1979, 1996); and hydrodynamic lubrication (Brodsky and Kanamori, 2001). While such processes may be operative during an earthquake, they do not relate to the problem of how earthquakes initiate, and each requires that very specific fault-zone conditions exist to be viable. Thus, to assess the likelihood that dynamic weakening mechanisms might operate along active faults, we need to compare the structure and physical properties of actual fault zones with parameters required by these various models and then combine these observations with the results of long-term fault-zone monitoring in the near-field of small-to-moderate size earthquakes.

4.07.4 Near-Field Observations of Earthquake Nucleation and Propagation

Understanding the physical processes operating during both nucleation and rupture propagation can prove to be critical to understanding why plate-boundary faults tend to be anomalously weak, especially if dynamic weakening mechanisms

are important. In addition, by drilling into active faults at seismogenic depths and observing repeating earthquakes at very short distances, one can observe near-field phenomena for earthquakes of $M \sim 1$ or larger, thereby providing a new window into the physics of the earthquake source. Ideally, in addition to instrumenting fault-zone drill holes with seismometers, it would be ideal to place instruments within or immediately adjacent to active sliding surfaces to directly monitor fault displacement, deformation, pore pressure, and heat generated during sliding. The work of Brune and coworkers on foam rubber models of earthquakes (Anooshehpoor and Brune, 1994; Brune et al., 1993) illustrates the advantages of making measurements at or very near the sliding surface. Many of the objectives for near-field observation would be met by sensors placed within a few hundred meters of the earthquake source. At these distances, near-field waves will be of significant amplitude compared to the far-field waves for $M \geq 0$ events, and static strains will be well within the resolution of borehole strainmeters.

The process by which the fault becomes unstable and initiates a dynamically propagating rupture is central to understanding how earthquakes work. It has recently been proposed that the very beginnings of rupture for earthquakes in the magnitude range from at least $M_w = 1$–8 characteristically involve a period of slow growth of the seismic moment (Beroza and Ellsworth, 1996; Ellsworth and Beroza, 1995, 1998; Iio, 1992, 1995). The characteristics of this process, called the seismic nucleation phase, rule out self-similar models for the nucleation and growth of rupture including the standard model of a dynamically growing crack (Kostrov, 1964). Although a range of hypotheses have been proposed to explain this slow beginning to earthquakes, far-field observations have thus far proved to be inadequate to determine if the seismic nucleation phase represents a cascade of smaller events, in which case the dynamically expanding crack model might apply, or if it represents a transition from an aseismic (stable) sliding to dynamic rupture, as required by laboratory-based and theoretical models of rupture initiation (Dieterich, 1992; Ohnaka, 1992). Observations of the nucleation process made within the near field in drill holes passing close to or through active faults at seismogenic depths have the potential to resolve this process, as they will not be distorted by attenuation or scattering, which limits the interpretation of available data (Iio, 1995).

The physics of earthquake rupture propagation has also been the subject of intensive investigation in recent years (e.g., Brune et al., 1993; Harris, 2004; Heaton, 1990; Melosh, 1996; Peyrat et al., 2004). New data have again drawn into question the standard model of a dynamically expanding crack that heals inward from its outer boundary (Madariaga, 1976). There is now evidence from large earthquakes that the rupture may propagate as a 'slip pulse' (e.g., Wald and Heaton, 1994), yet we know little about how such a concentrated slip zone is generated or maintained or why the fault comes to rest so abruptly. Brune et al. (1993) had further proposed that tensile opening of the fault accompanies the shear displacement in the slip pulse. If correct, it would be a mechanism by which a fault can have high static strength but slide without generating heat. However, sliding at near-zero normal stress implies that the dynamic stress drop should equal the tectonic stress (see Lachenbruch and Sass, 1980) resulting in near-zero shear stress on the fault after rupture (e.g., Zoback and Beroza (1993) for the Loma Prieta earthquake). Thus, measuring the dynamic stress drop in the near-field region will give us a direct test of the high static strength/low dynamic friction hypothesis.

Recent observations of microearthquakes in moderately deep boreholes (2–2.5 km) at Cajon Pass, California (Abercrombie, 1995; Abercrombie and Leary, 1993), at the SAFOD pilot hole (Imanishi et al., 2004), and in Long Valley, California (Prejean and Ellsworth, 2001) demonstrate that ultrahigh-fidelity recordings can be made in downhole observatories. Imanishi and Ellsworth (2006) recently used high-frequency recordings from the SAFOD pilot hole array to study the scaling of apparent stress (seismic wave energy/seismic moment). They found that this ratio shares the same upper bound over the magnitude range from M 9 to M 0, providing strong support for the self-similar model of the earthquake source.

4.07.5 Fault-Zone Drilling Projects

In this section, we review several recent, ongoing, and planned scientific drilling projects related to earthquake studies. Many of the key questions to be addressed by deep drilling into active faults were enumerated in December 1992 at a workshop (attended by 113 scientists and engineers from seven countries) on San Andreas Fault zone drilling at the Asilomar Conference Center in Pacific Grove, California. While the purpose of this workshop was to initiate a broad-based scientific discussion of the issues that could be addressed by drilling and direct experimentation in the San Andreas Fault, the arguments put forth for scientific drilling into fault zones are generally applicable to all of the fault-zone drilling projects discussed in this section.

The questions to be addressed by fault-zone drilling projects are far-reaching and include the following:

4.07.5.1 Fault Behavior

- What is the static strength of active faults? Why are some faults anomalously weak?
- Why are some faults (or fault segments) creeping and some locked?
- What factors control the localization of slip and strain?
- How are faults stressed at different crustal levels?
- How does strain communication occur within fault zones over different timescales?
- How is energy partitioned within fault zones between seismic radiation, frictional dissipation, grain size reduction, and chemical reactions?
- Can the frequency–magnitude relationship for earthquakes be extrapolated to smaller magnitudes?

4.07.5.2 Fluid Pressure

- What is the vertical and lateral distribution of fluid pressure regimes in faults?
- Do fault fluid pressure compartments exist?
- If so, what is the nature of the seals between these compartments?

- What is the time dependence of fluid pressure within fault zones?
- What is extent of vertical and lateral fluid migration during a seismic stress cycle?

4.07.5.3 Fault Fluids

- What is the origin and composition of fault-zone fluids?
- What are the permeabilities of fault-zone materials and country rock?
- What are the fluid transport mechanisms in and adjacent to fault zones and what physical processes lead to fluid redistribution?
- What is the interplay between water–rock interaction and rheology at different structural levels?

4.07.5.4 Fault-Zone Properties and Physical Parameters

- How does the stress tensor vary in the vicinity of fault zones?
- How do pre- and postfailure stress states compare?
- What, if any, form of cyclical dilatancy operates in the vicinity of fault zones?
- How do physical properties relate to the fault-zone fabric?
- What is the origin of low-velocity zones associated with fault zones?
- How well and in what manner do physical properties and heterogeneity measured in boreholes correlate with geophysical observables?

4.07.5.5 Fault Structure and Materials

- How does the width and character of active slip zones vary with depth?
- What is the thermal structure of active fault zones?
- How do mineralogy and deformation mechanisms within fault zones change with depth, temperature, and country-rock geology?
- What determines the maximum depth of seismic activity?
- At what temperature do mineral reaction kinetics operate at the timescale of an earthquake cycle?
- How accurate are inferences drawn from deformation microstructures, piezometers, and fluid inclusions and how might one assess their survivability?

Fundamental questions about faulting and earthquakes such as these have gone unanswered due to the complete lack of data on the physical and chemical processes operating on faults at depth. Hence, the principal reasons for drilling into active faults are to conduct extensive investigations in situ and on exhumed materials that are representative of the faults at the pressures, temperatures, and conditions at which major earthquakes nucleate.

4.07.5.6 San Andreas Fault Observatory at Depth

The site identified for this 3.0 km deep drilling project is located in Central California (Figure 2), about 10 km northwest of the town of Parkfield. This site is located at the northwestern end of the rupture zone of the 1966 and 2004 $M=6$ Parkfield earthquakes, in the transition between the creeping and locked sections of the San Andreas Fault. The San Andreas displays a range of behaviors at this site. At the surface, the fault is creeping at a rate of 1.8 cm year^{-1}, with most of the fault displacement localized to a zone no more than 10 m wide (Burford and Harsh, 1980; Schulz, 1989). Numerous earthquakes occur directly on the San Andreas Fault in the depth interval from about 3 to 12 km. The shallow seismicity at Parkfield occurs in tight clusters of activity (Nadeau et al., 1994, 1995, 2004) that have remained spatially stationary for at least the past 20 years.

An important feature of the microearthquakes beneath Middle Mountain is that they occur in families of repeating events, which provide a reliable target for guiding drilling (Figure 3). Individual earthquakes have been observed to recur numerous times using the UC Berkeley High-Resolution Seismic Network, at precisely the same location and with the same magnitude (Nadeau and McEvilly, 1997; Nadeau et al., 1994, 1995, 2004). Repeating sources of up to $M=2$ are located at drillable depths beneath the proposed drill site. A major goal of this experiment is to drill as close as possible to one or more of these sources and to follow the buildup of strain and its release through multiple earthquake cycles during the monitoring phase of the experiment.

Illustrated in Figure 3 is a 2 km deep pilot hole drilled at the SAFOD site in 2002, 1.8 km to the southwest of the surface trace of the San Andreas Fault (see review by Hickman et al. (2004) and associated papers published in special issues of Geophysical Research Letters). During the summers of 2004 and 2005, the

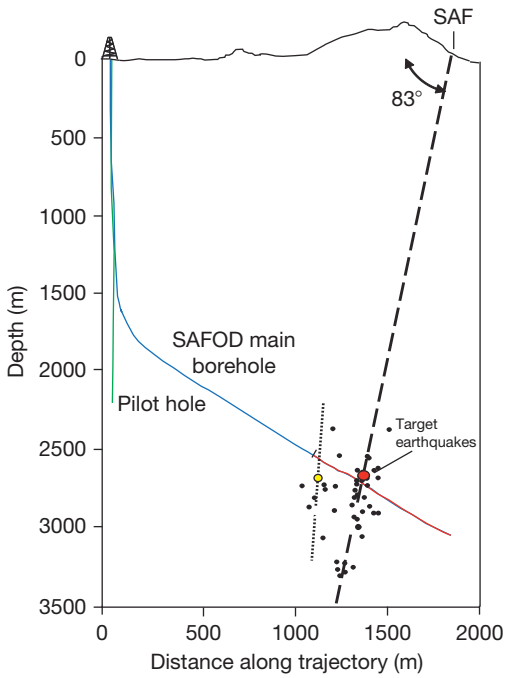


Figure 3 Schematic diagram of the SAFOD pilot hole and main hole. The precise location of the target earthquakes is approximate although the relative locations are exact. As indicated by the distribution of microearthquakes, slip is accommodated by multiple faults at depth. Simply connecting the position of the San Andreas Fault at depth with the surface position implies that the fault dips steeply to the west.

SAFOD main hole was rotary drilled through the zone of the repeating microearthquakes. Note that in the subsurface, the position of the fault is to the southwest of the surface trace.

Four major geologic units were encountered along the trajectory of the SAFOD main hole: In the vertical section of the wellbore, the near-surface Quaternary and Tertiary sediments were found to be underlain by Salinian granite at a depth of ~700 m. After deviating the borehole toward the fault, arkosic sediments (most likely locally derived from Salinian granite) were encountered about 300 m NE of the drillsite, perhaps after crossing the Buzzard Canyon fault, a NW-trending strike-slip fault exposed at the surface that trends subparallel to the San Andreas (M. Rymer, personal communications). Approximately, 1200 m NE of the drillsite, a possibly ancestral trace of the San Andreas was crossed as the lithology changed abruptly to claystones and siltstones of the Great Valley formation, found throughout Central California on the east side of the San Andreas.

Geophysical logs and cuttings analyses indicate that the San Andreas is a broad zone of anomalously low *P*- and *S*-wave velocity and resistivity that define a relatively broad damage zone as indicated in Figure 4. The locations of active fault traces are revealed by casing deformation (indicated by the red line in Figure 4) and the location of a $M \sim 0$ microearthquake that occurred in May 2005 (indicated in Figure 4). Note that these active faults are associated with narrower, more highly localized zones of low *P*- and *S*-wave velocity and resistivity embedded within the broader damage zone. In a third phase of the project, coring will be done in multilateral holes to directly sample the damage zone and both creeping and seismically active faults at depth. Preliminary results from the project were presented in several special sessions of the December 2005 American Geophysical Union Fall Meeting (published in EOS, vol. 86, 2005), and numerous publications are in print or under review detailing early results from SAFOD.

An array of downhole seismometers, accelerometers, and other sensors will be deployed near the fault zone after drilling. This will allow for accurate determinations of the radiated energy, seismic moment, and earthquake locations to be made, the detailed velocity structure of the fault zone to be investigated, and earthquake nucleation and propagation to be studied with unprecedented detail. High-frequency signals are expected from the very local events, with good signal-to-noise above 1 kHz. Thus, good sensitivity to the higher frequencies becomes the most important seismic monitoring design criteria. Peak acceleration estimates for a magnitude 1.0 earthquake located at a distance of 1000 and 100 m are 1 and 10 g, respectively.

4.07.5.7 Nojima Fault

In 1995, a *M* 6.9 strike-slip faulting earthquake struck Kobe, Japan, resulting in 6432 fatalities. The earthquake occurred on

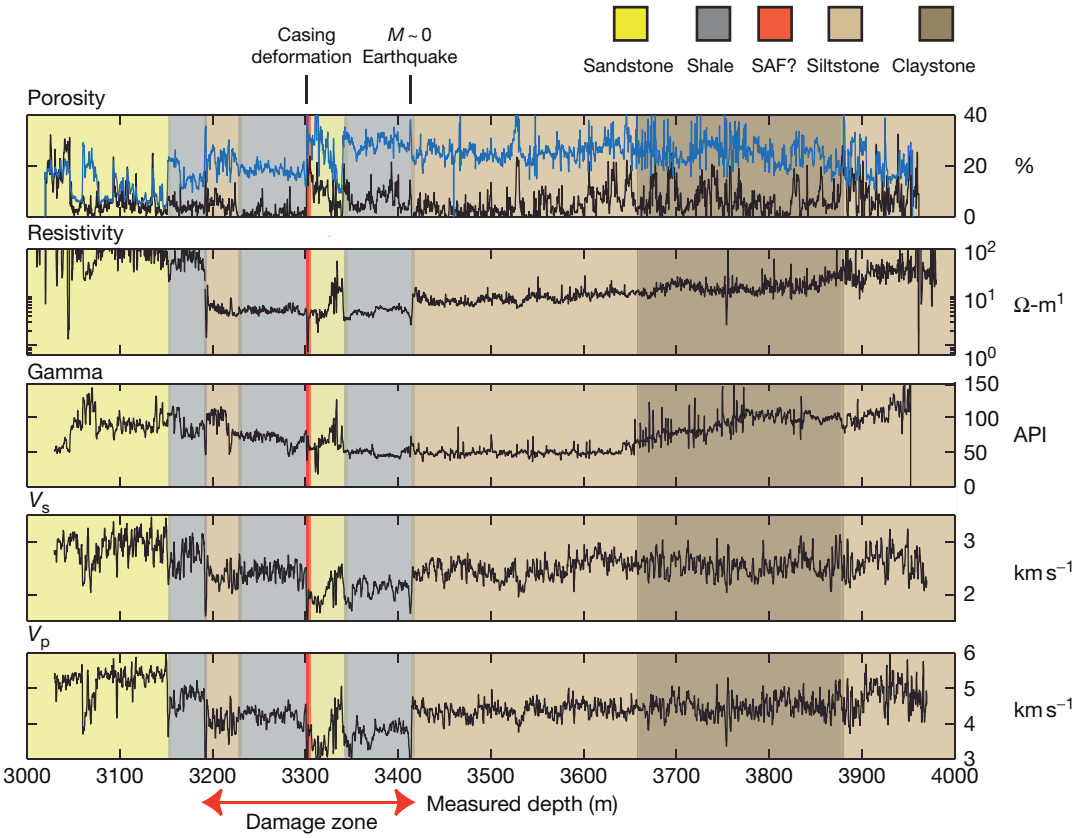


Figure 4 Geophysical logs and generalized lithology from phase 2 of the SAFOD project (shown in red in Figure 3). Anomalously low V_p, V_s, and resistivity indicate a broad low damage zone ~250 m wide. Multiple active fault traces appear to occur within the damage zone. The red line at 3300 m depth indicates an area of localized casing deformation. A $M \sim 0$ earthquake was recorded by a seismic array in the borehole that appears to be associated with the localized low-velocity zone at 3410 m.

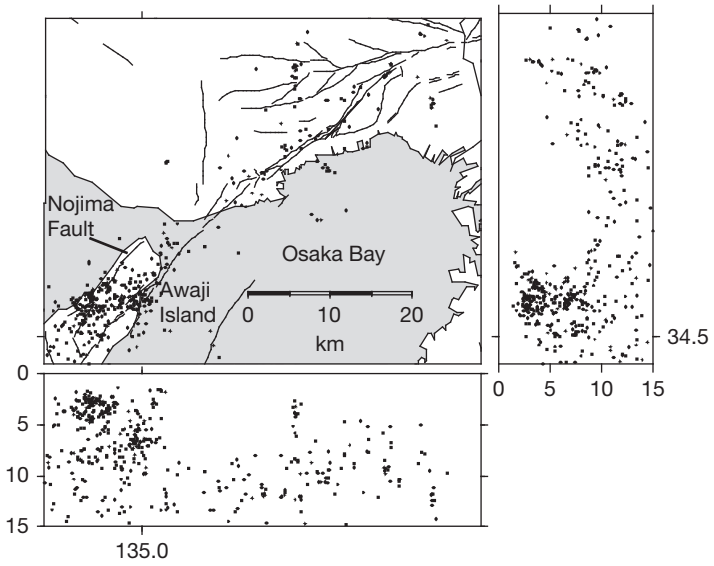


Figure 5 The Nojima Fault – source of the *M* 6.9 1995 Kobe earthquake – is an NE-trending strike-slip fault exposed on Awaji Island and extends beneath Osaka Bay. Numerous shallow aftershocks (black dots) occurred on the section of the fault where drilling took place. East–west and north–south cross sections show the distribution of aftershocks with depth. As shown, many of the aftershocks beneath Awaji Island are extremely shallow (2–5 km deep), with others extending to typical depths of 10–15 km. Courtesy of H. Ito.

the Nojima Fault (Figure 5), a right-lateral strike-slip fault that trends NE–SW through Awaji Island and beneath Osaka Bay (Ando, 2001). Several research groups drilled boreholes (ranging in depth from 747 to 1800 m) through the fault along segments of the Nojima Fault that exhibited surface rupture during the earthquake. Note the numerous shallow (2–5 km depth) occurred on the Nojima Fault beneath Awaji Island. Extensive geophysical logging, detailed core analysis, stress measurements, and repeated hydrologic tests and repeated injection tests were carried out. The preliminary results were published in the special volume of the Island Arc (Oshiman et al., 2001).

Two boreholes were drilled at Hirabayashi (where the maximum slip was 2 m) that penetrated the core of the Nojima Fault. The structure of the Nojima Fault zone is characterized by a narrow fault core with various types of fault gouge. The hanging wall of the fault displays many minor shear zones that increase in frequency toward the fault core (Figure 6). The footwall of the fault is significantly less deformed and less altered with respect to the hanging wall. The physical properties of fault zone and country rocks have been extensively studied (Boullier et al., 2001; Fujimoto et al., 2001; Ohtani et al., 2001; Tanaka et al., 2001).

Integrated borehole monitoring systems (three-component strain meter, three-component seismometer, and water-level sensors) have been installed along two different segments of the Nojima Fault.

4.07.5.8 Chelungpu Fault

The 1999 *M* 7.7 Chi-Chi earthquake in Taiwan produced large zones of surface rupture, with a maximum displacement of 8 m on the Chelungpu Fault. Along nearly all of its length, the earthquake fault strikes ~N–S (Figure 7(a)) with the rocks to the east thrust over those to the west along an east-dipping thrust faults (Figure 7(b)). The northern portion of the fault is characterized by very large slip but relatively low ground acceleration during the *M* 7.7 earthquake. The TCDP was carried out near the town of Dakeng in a zone of large surface offset (and subsurface slip) near the northern end of the N–S-trending rupture surface.

The main scientific objectives for drilling were related to testing various mechanisms of nucleation and rupture of large earthquakes and provide answers to fundamental questions pertaining to the relationship between the Chelungpu thrust and regional tectonics (Ma et al., 2003; Tanaka et al., 2002). TCDP drilling was started in 2004, and two boreholes were completed in 2005, including successful continuous coring, logging, and borehole monitoring. The direction of maximum horizontal compression found in one of the TCDP boreholes is ~N 120° E, consistent with the regional stress field (Kao and Angelier, 2001) and relative motion of the Philippine Sea Plate with respect to Taiwan (Wu et al., 2007).

As shown in Figure 8, two holes (A and B) were drilled through the main fault zone that ruptured during the Chi-Chi earthquake (Figure 8). In hole A, multiple fault zones were identified in the Pliocene Chinshui Shale and Miocene Kueichulin Formation, but the primary slip zone in hole A was encountered at a depth of 1111 m (see photo in Figure 8) and 1136 m in hole B (see photo in Figure 8). This fault is associated with bedding-parallel thrusting with a gentle dip of about 20° and is characterized by over 1 m thickness of gouge (fault core) and gradational breccia in a damage zone found in both the upper and lower blocks. Wu et al. (2007) noted a major perturbation of the stress field in hole A in the vicinity of the Chelungpu Fault.

To address key questions of earthquake rupture dynamics such as energy dissipation and fracture energy, Ma et al. (2006) analyzed the grain size distribution of the fault gouge. They estimated that 6% of the total earthquake energy is associated

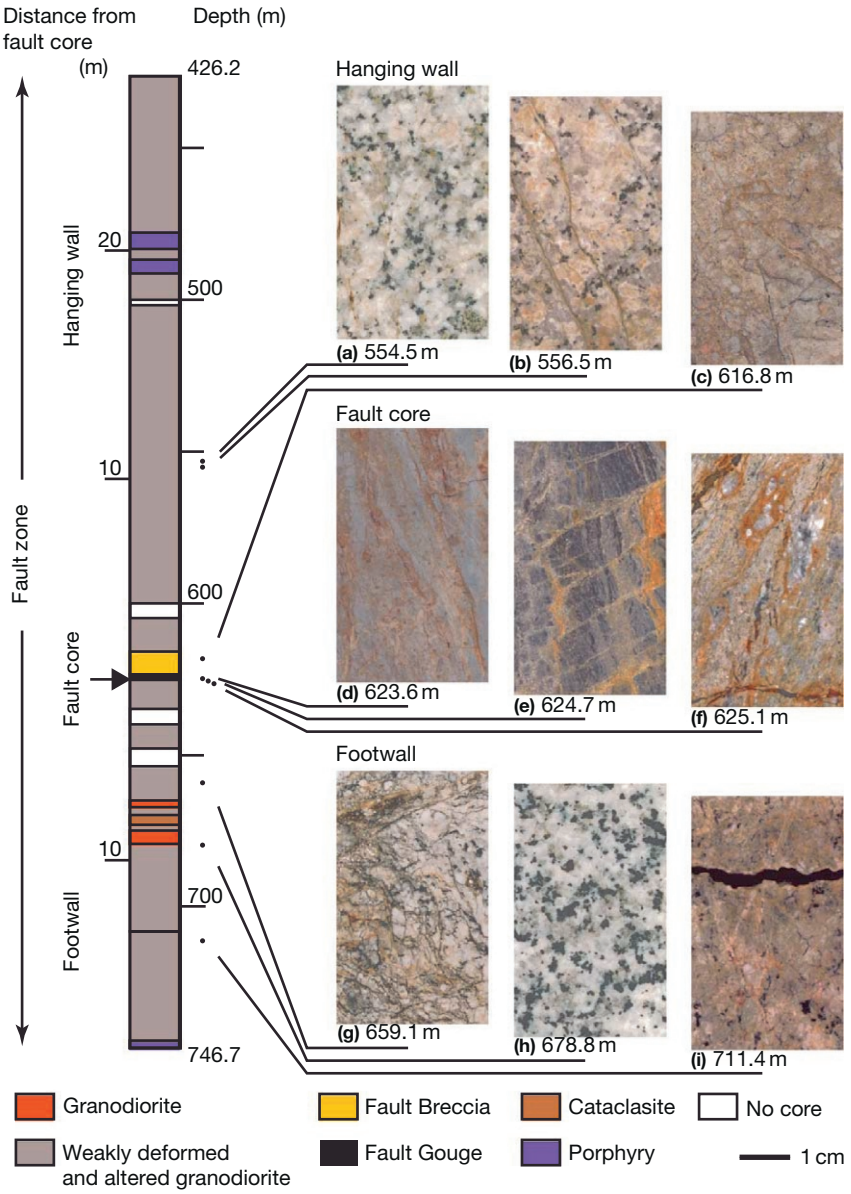


Figure 6 Lithologies encountered in the Nojima Fault and photographs of rocks from the hanging wall, fault core, and footwall. Reproduced from Ando M (2001) Geological and geophysical studies of the Nojima Fault from drilling: An outline of the Nojima Fault Zone Probe. *The Island Arc* 10: 206–214.

with gouge formation. Tanaka et al. (2007) and Kano et al. (2006) discussed the possibility that a thermal anomaly measured at the trace of the fault is associated with frictional heat generated during faulting.

4.07.5.9 Gulf of Corinth

The Gulf of Corinth, Greece, is the most actively deforming region in Europe and one of the most seismically active areas. Deformation in the area is dominated by north–south-directed back-arc extension associated with the Hellenic Trench. The Gulf of Corinth is an east–west-trending graben bordered on the north and south by east–west-trending normal faults (Figure 9). The extension rate in the region ranges between 1 and 1.5 cm year^{-1}. A number of large earthquakes have occurred in the region in the past century.

The Corinth Rift Laboratory (CRL) project involved drilling into the Aigion Fault on the south side of the Gulf. This site offers excellent conditions for in situ investigation of active normal faulting, the rifting process, and monitoring fluid–fault interactions (Cornet et al., 2004).

The AIG10 borehole was drilled in 2002–1000 m depth and intersected the Aigion Fault at 760 m (Figure 10). A key goal of the laboratory is to understand the relationships between outcrops of steeply dipping fault zones and the seismogenic faults at great depth. Special attention is directed to the interactions between circulating fluids and fault mechanics, including hydro-thermomechanical coupling and the role of geochemical healing and alteration. After passing through an overburden of sands, conglomerates, and clays, the borehole penetrated alternating Mesozoic radiolarite and limestone in the hanging wall before intersecting the main fault zone at a

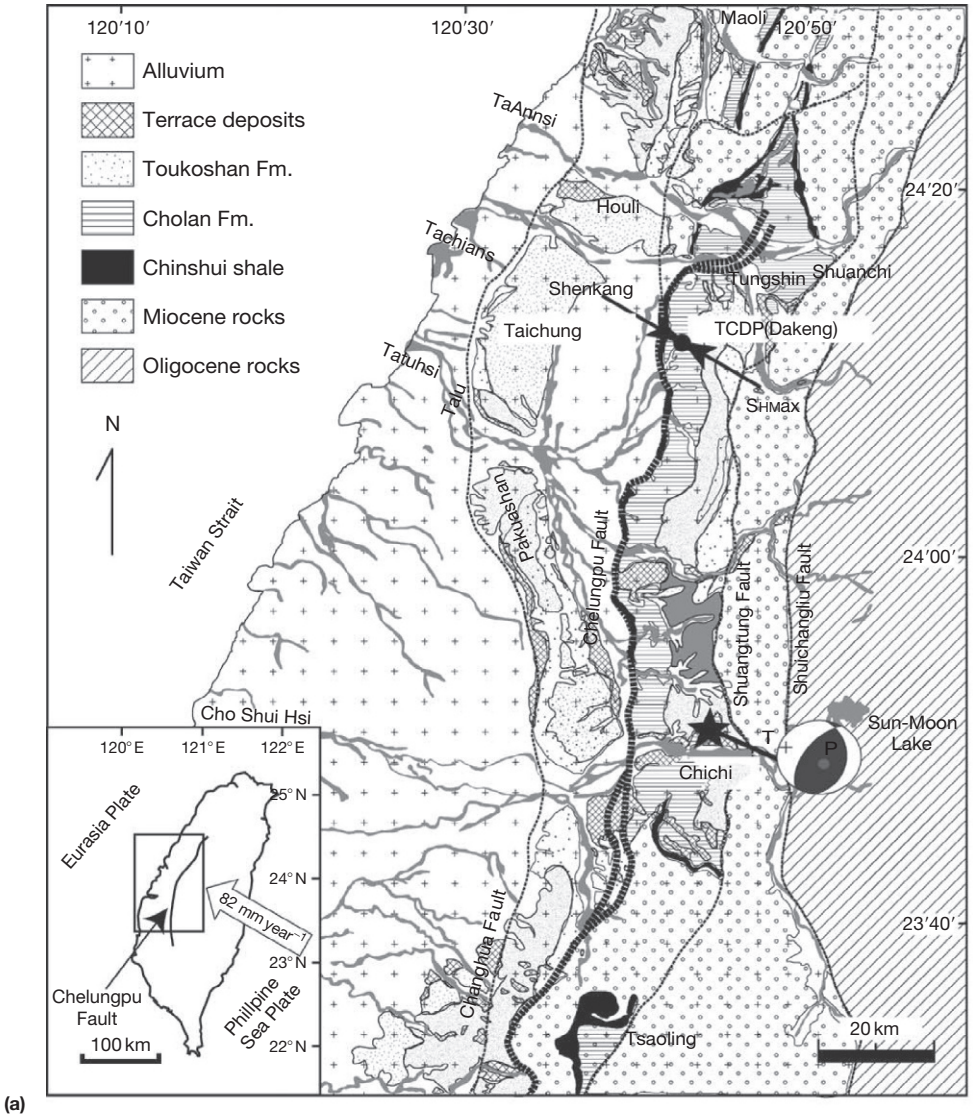


Figure 7 (a) Generalized map of surface geology in western Taiwan (H. Tanaka, personal communication, 2006): the trace of the Chelungpu Fault that slipped in the *M*7.71999 Chi-Chi earthquake epicenter and the location of the TCDP boreholes near Dakeng are shown. The average ~N 120° E direction of maximum horizontal compression observed in the borehole (Wu et al., 2007) is the same as the regional stress field (Kao and Angelier, 2001) and the motion direction of the Philippine Sea Plate relative to Taiwan (see inset). (b) Geologic cross section derived from high-resolution seismic reflection data showing the location of the TCDP borehole with respect to the east-dipping Chelungpu Fault. Reproduced from Yue LF, Suppe J, and Hung JH (2005) Structural geology of a classic thrust belt earthquake: The 1999 Chi-Chi earthquake Taiwan (*M*7.6). *Journal of Structural Geology* 27: 2058–2083.

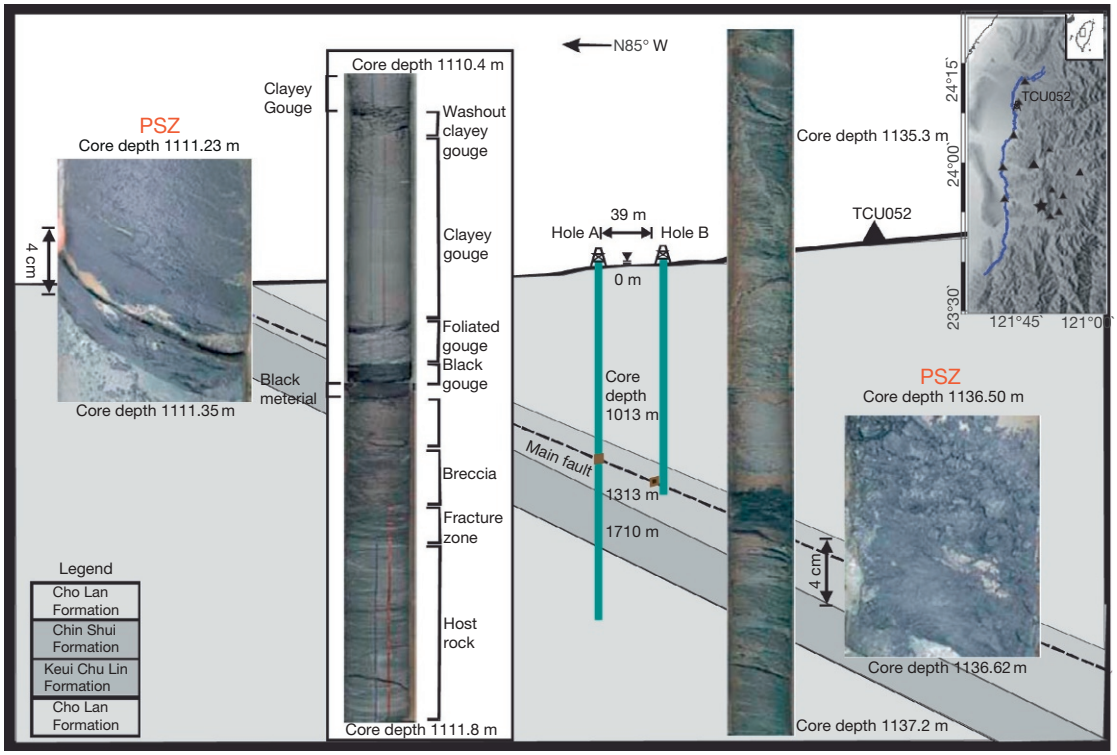


Figure 8 Detailed lithologies and fault rocks encountered in the two TCDP core holes shown on a background of the generalized geology as shown in Figure 7(b). The primary slip zone in hole A was encountered at a depth of 1112.23–1111.35 m (see photo) and 1136.50–1136.62 m in hole B (see photo). Courtesy of K.-F. Ma.

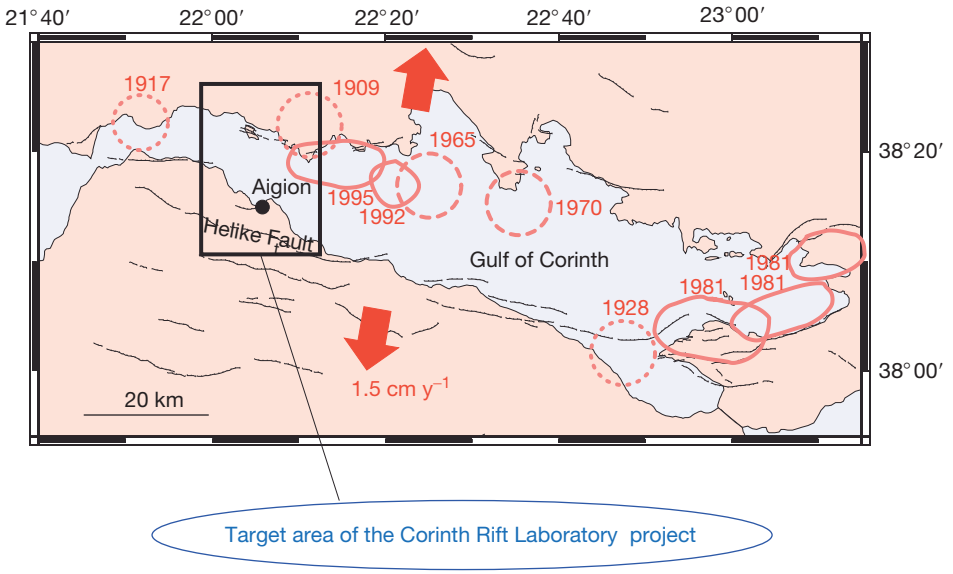


Figure 9 Active tectonics of the Gulf of Corinth area showing the locations of active normal faults, large earthquakes of the past 100 years, and geodetically determined extension directions and rates. Courtesy of F. Cornet.

depth of 760 m, where cataclastic fault rocks were encountered. Below the fault, a homogeneous, heavily karstified limestone was encountered down to 1000 m.

The Aigion Fault zone forms a hydraulic barrier that sustains about a 0.5 MPa differential pressure across it. There is a slight overpressure (~0.5 MPa) in the hanging wall immediately above the fault with a larger overpressure (0.9 ± 0.1 MPa) in the footwall resulting from the topographic relief from the footwall to the hanging wall sides of the fault. Nearly isothermal temperatures (31 °C) below the fault are consistent with local convection in the karstic limestone, while there is an absence of flow above the fault. Geochemical

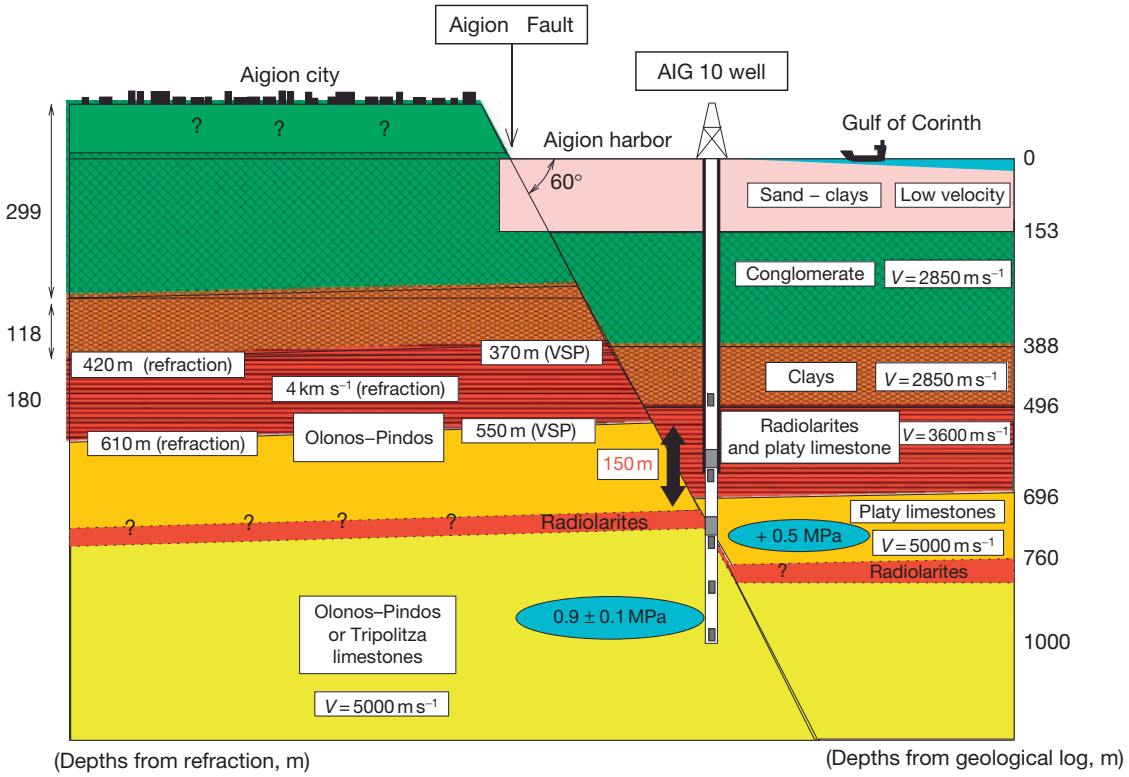


Figure 10 Geology in the vicinity of the AIG 10 well that was cored through the Aigion Fault. The fault was encountered at a depth of ~760 m. Pore pressure on the footwall side of the fault at depth was ~1 MPa in excess of hydrostatic, consistent with the fact that the footwall elevation is ~100 m higher than the elevation of the well. Courtesy of F. Cornet.

data indicate a shallow continental origin of water in and adjacent to the fault zone, with a notable absence of deep fluid input from the mantle. The borehole is being used to make continuous measurements of pressure in the fault zone.

The data acquired to date show tidal variations, as well as pressure variations induced by remote earthquakes (Cornet et al., 2004). The Corinth science team proposes to drill a 4.5 km deep borehole to monitor transients in pore pressure within the seismogenic zone and provide clues to the origin of deep fluids.

4.07.5.10 NELSAM

The deep gold mines of South Africa offer unique environments to study earthquakes by providing access to the focal area. The mining operations generate thousands of earthquakes per day, and some of these events approach *M*5. The Natural Earthquake Laboratory in South African Mines (NELSAM) is a project that utilizes the availability of these deep mines for earthquake research (Reches et al., 2006). The central part of this project is dense instrumentation and detailed characterization of a large fault zone in TauTona Mine, which is the deepest mine on Earth (approaching 4 km, depth) located within the Western Deep Levels of the Witwatersrand Basin, South Africa (Figure 11).

The laboratory is built around the Pretorius fault zone (shown in pink in the lower part of Figure 11), which is at least 10 km long with 30–200 m of displacement and which appears to have been inactive for the past 2.5 Ga. The mining plan for the next few years is likely to induce earthquakes of significant magnitude ($M > 2$) along this fault. NELSAM science activities at the site started in 2005 with site characterization, including mapping of 3-D structure and composition of the Pretorius fault zones at 3.54 km depth. The emphasis of the investigations has been on segments of the fault zone that were reactivated during recent earthquakes. Analysis of the stress state both near and away from the excavation will be carried out using techniques developed for the oil industry in which compressive and tensile wellbore failures are studied in wells of different orientation (e.g., Zoback et al., 2003). In addition, five boreholes, 20–60 m in length, have been drilled across the Pretorius fault zone. Once completed, the earthquake laboratory will include a dense array (250 m footprint) of three-component broadband accelerometers, seismometers, strain meters, temperature sensors, creep meters, electromagnetic radiation sensors, and acoustic emissions sensors. Fault-zone fluid chemistry will be monitored with an onsite mass spectrometer.

A *M* 2.2 earthquake that occurred on 12 December 2004 in the center of the planned earthquake laboratory reactivated several segments of the Pretorius fault (Heesakkers et al., 2005). The mapping of the rupture zone of this earthquake in three dimensions revealed quasi-planar, crosscutting reactivated segments with inclinations ranging from 21° to 90°. The rupturing formed fresh fine-grained white rock powder almost exclusively along the contacts of the ancient, sintered cataclasite and the quartzitic host rock.

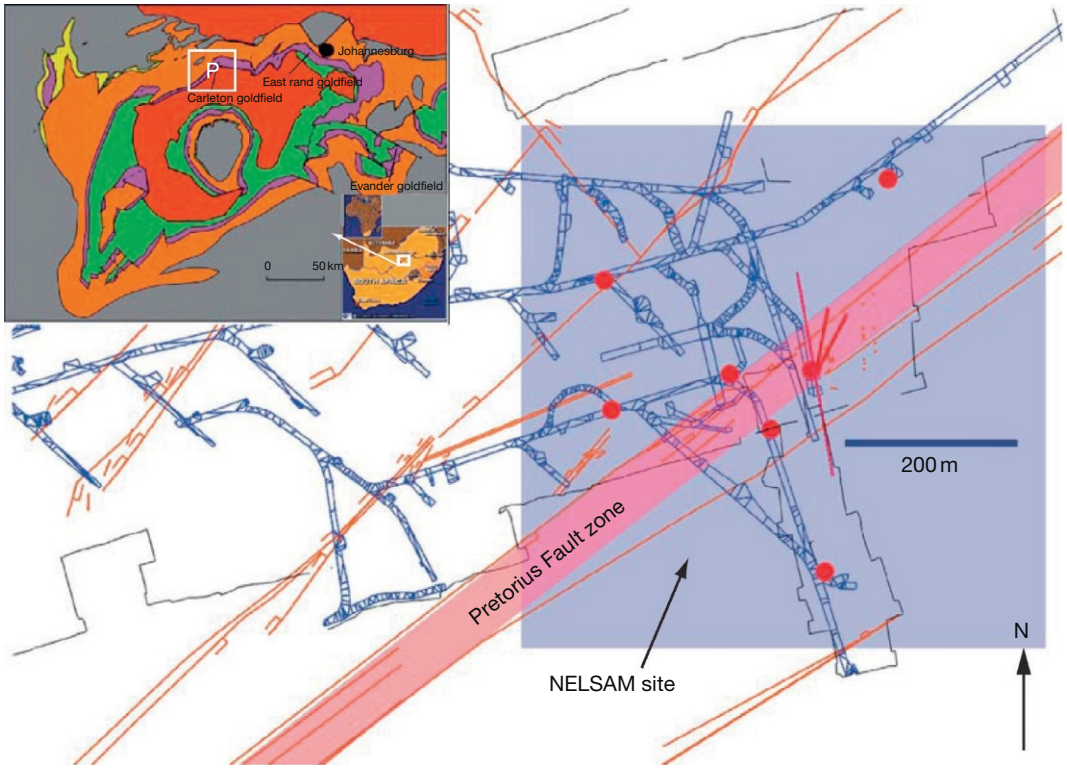


Figure 11 The TauTona Mine is located in the Carleton goldfield, ~80 km west of Johannesburg, South Africa (see inset). Scientific drilling is occurring in the vicinity of the Pretorius fault zone (shown by pink shading) where a number of recent earthquakes have occurred recently. The figure shows the excavations at a depth of 3597–3657 m. Most of the faults exposed in the mine tunnels trend NE–SW, similar to the overall trend of the Pretorius fault zone. Courtesy of Z. Reches.

4.07.5.11 NanTroSeize

An extensive fault-zone drilling campaign is planed as part of NanTroSeize experiment (Figure 12(a)) to be carried out by the Integrated Ocean Drilling Program in the Nankai Trough just off southern Honshu, Japan. The extent of the 1944 Tonankai earthquake slip zone, the presumed extent of the locked zone (indicated by the heavy white line), and the presumed seismogenic zone are also shown.

The first-order objectives are to

- document the material properties and state of the plate-boundary fault system at several *P*–*T* and lithology conditions, testing hypotheses for stable versus unstable frictional behavior;
- investigate partitioning between seismic and aseismic processes on the main plate boundary, through monitoring of seismicity, borehole strain, and pore fluid pressure;
- test whether there are interseismic temporal changes in state along the plate-boundary thrust, including possible earthquake precursory signals; and
- use downhole measurements and sampling to calibrate observations made in the broader geophysical volume surrounding the boreholes.

In many ways, the specific science goals of this experiment are somewhat similar to those of the SAFOD project, although in a much different geologic setting. Specifically, the goals are to

- obtain samples of faults and surrounding environment; characterize lithology, structure, elastic and mechanical properties, porosity, permeability, pore fluid chemistry, and microbiology; and use familiar ODP coring techniques, augmented by cuttings and sidewall coring;
- characterize the near-borehole environment, using geophysical logging, wireline and logging while drilling, and active testing for pore fluid pressure, stress, and hydrogeologic properties (permeability and storage);
- monitor the borehole environment over time, passive and active source seismology, strain and tilt, pore fluid pressure, temperature, and EM field.

A detailed view of the proposed drilling area is shown in Figure 12(b) (see box in Figure 12(a)) where seismic reflection data reveal the detailed structure of the main décollement and splay faults in the accretionary prism. As shown, a series of relatively shallow holes will be drilled through splay faults in the accretionary wedge to study aseismic faulting. Following drilling, these holes will be used for downhole observatories. An eventual ultradeep borehole is to be drilled into a seismogenic section of the décollement and possibly seismogenic deep part of the splay fault.

4.07.6 Summary

Drilling and downhole measurements in active fault zones provide critical tests of hypotheses arising from seismological observations, laboratory rock deformation experiments, and geologic observations of exhumed fault zones. Drilling provides the only direct means of measuring pore pressure, stress,

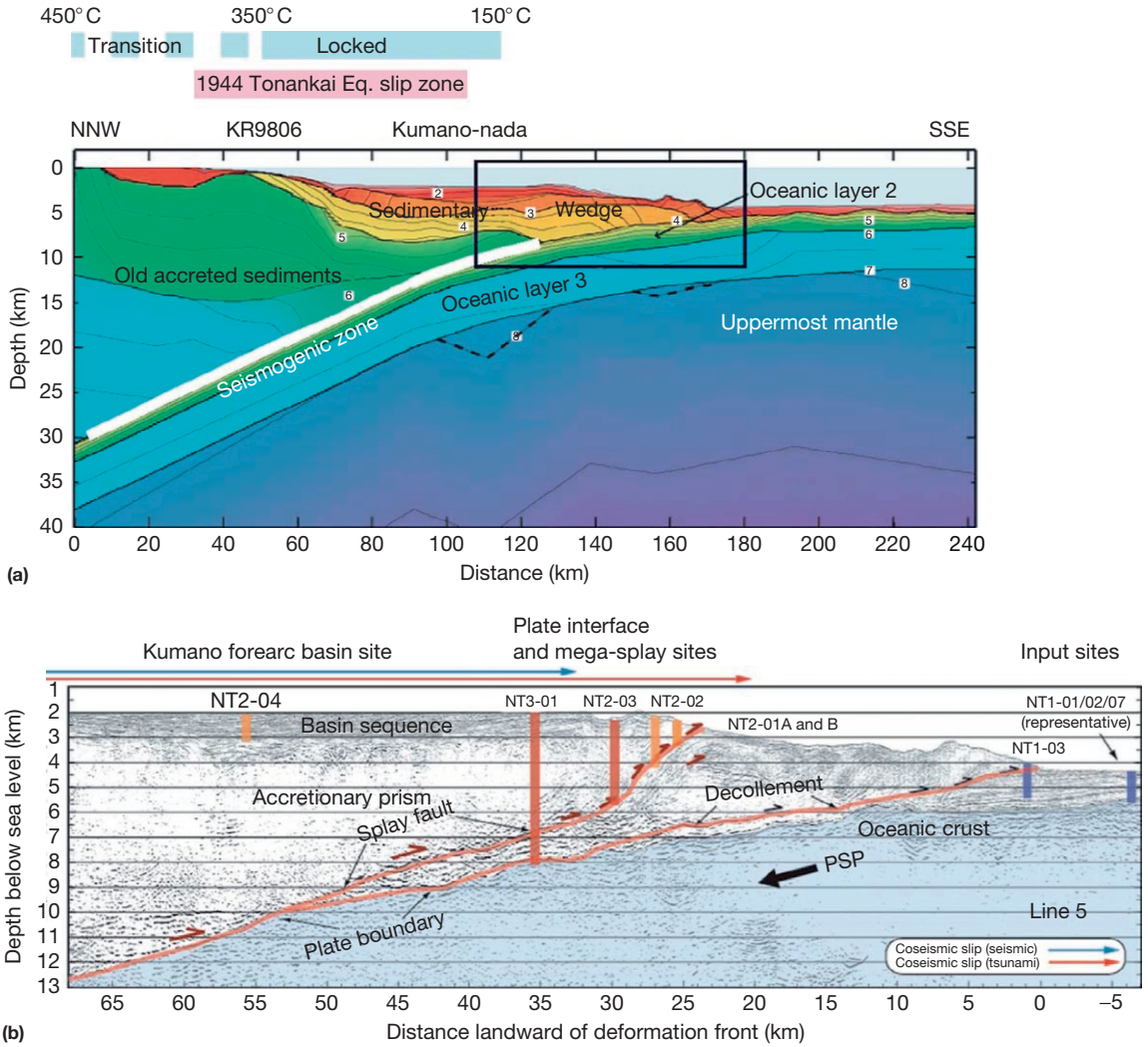


Figure 12 (a) Seismic velocity structure in the Nankai Trough area reveals the nature of the downgoing Philippine Sea Plate, the accretionary wedge, and old accreted sediments (reproduced from Nakanishi A, Takahashi N, and Park J-O, et al. (2002) Crustal structure across the coseismic rupture zone of the 1944 Tonankai earthquake, the central Nankai Trough seismogenic zone. *Geophysical Research* 107: http://dx.doi.org/10.1029/2001JB00424). The extent of the 1944 Tonankai earthquake slip zone, the presumed extent of the locked zone (indicated by the heavy white line), and the presumed seismogenic zone are also shown. (b) A detailed view of the proposed drilling area (see box in na) where seismic reflection data reveal the detailed structure of the main décollement and splay faults in the accretionary prism. Note that a series of boreholes are planned to penetrate various parts of the main décollement and splay faults. Courtesy of H. Tobin.

permeability, and other important parameters within and near an active fault zone at depth. It is also the only way to collect fluid and rock samples from the fault zone and wall rocks at seismogenic depths and to monitor time-dependent changes in fluid pressure, fluid chemistry, deformation, temperature, and electromagnetic properties at depth during the earthquake cycle. In the context of the processes and properties alluded to in the preceding text, in situ observations and sampling through drilling perform two critical, and unique, functions. Sampling of fault rocks and fluids and downhole measurements provide essential constraints on mineralogy, grain size, fluid chemistry, temperature, stress, pore geometry, and other parameters that would allow laboratory investigations of fault-zone rheology and frictional behavior to be conducted under realistic in situ conditions. By in situ sampling, downhole measurement, and long-term monitoring in active fault zones, one is able to test and refine the broad range of current theoretical models for faulting and seismogenesis by providing realistic constraints on fault-zone physical properties, loading conditions, and mechanical behavior at depth. In particular, by comparing results of microstructural observations and rheological investigations on core with measurements of microseismicity, fluid pressure, and deformation during the fault-zone monitoring phase of this experiment, we would be able to differentiate between fault-zone processes (e.g., fluid pressure fluctuations) associated with fault creep versus earthquakes.

Of course, no single drill hole into an active fault can address all of the questions listed earlier. Rather, like other avenues of earthquake research, the experiments being carried out in fault-related scientific drilling projects around the world (such as the ones noted earlier) cumulatively address many critical scientific questions about fault-zone structure, composition, and processes. Over time, the accumulation of data from these experiments will result in important advances in earthquake science.

References

Abercrombie RE (1995) Earthquake source scaling relationships from −1 to 5 M_1 using seismograms recorded at 2.5 km depth. *Journal of Geophysical Research* 100: 24015–24036.

Abercrombie RE and Leary P (1993) Source parameters of small earthquakes recorded at 2.5 km depth, Cajon pass, Southern California: Implications for earthquake scaling. *Geophysical Research Letters* 20: 1511–1514.

Ando M (2001) Geological and geophysical studies of the Nojima Fault from drilling: An outline of the Nojima Fault Zone Probe. *The Island Arc* 10: 206–214.

Andreani M, Boullier A-M, and Gratier J-P (2005) Development of schistosity by dissolution–crystallization in a California serpentinite gouge. *Journal of Structural Geology* 27: 2256–2267.

Andrews D (2002) A fluid constitutive relation accounting for thermal pressurization of pore fluid. *Journal of Geophysical Research* 107(B12): 2363. http://dx.doi.org/10.1029/2002JB001942.

Andrews DJ and Ben-Zion Y (1997) Wrinkle-like slip pulse on a fault between different materials. *Journal of Geophysical Research* 102: 553–571.

Angevine CL, Turcotte DL, and Furnish MD (1982) Pressure solution lithification as a mechanism for the stick–slip behavior of faults. *Tectonics* 1: 151–160.

Anooshehpoor A and Brune JN (1994) Frictional heat generation and seismic radiation in a foam rubber model of earthquakes. *Pure and Applied Geophysics* 142: 735–748.

Bangs NL, Shipley TH, Gulick SPS, Moore G, Kuromoto S, and Nakamura Y (2004) Evolution of the Nankai Trough decollement from the trench into the seismogenic zone: Inferences from three-dimensional seismic reflection imaging. *Geology* 32: 273–276.

Barton CA, Zoback MD, and Burns KL (1988) *In-situ* stress orientation and magnitude at the Fenton Geothermal site, New Mexico, determined from wellbore breakouts. *Geophysical Research Letters* 15: 467–470.

Baumgärtner J and Zoback MD (1989) Interpretation of hydraulic fracturing pressure–time records using interactive analysis methods. *International Journal of Rock Mechanics and Mining Sciences and Geomechanical Abstracts* 26: 461–470.

Ben-Avraham Z and Zoback MD (1992) Transform-normal extension and asymmetric basins: An alternative to pull-apart models. *Geology* 20: 423–426.

Beroza GC and Ellsworth WL (1996) Properties of the seismic nucleation phase. *Tectonophysics* 261: 209–227.

Bird P (1984) Hydration-phase diagrams and friction of montmorillonite under laboratory and geologic conditions, with implications for shale compaction, slope stability, and strength of fault gouge. *Tectonophysics* 107: 235–260.

Blanpied ML, Lockner DA, and Byerlee JD (1991) Fault stability inferred from granite sliding experiments at hydrothermal conditions. *Geophysical Research Letters* 18: 609–612.

Boness N and Zoback MD (2006) A multi-scale study of the mechanisms controlling shear velocity anisotropy in the San Andreas Fault Observatory at Depth. *Geophysics* 7(5): F131–F146.

Bos B and Spiers CJ (2001) Experimental investigation into the microstructural and mechanical evolution of phyllosilicate-bearing fault rock under conditions favouring pressure solution. *Journal of Structural Geology* 23: 1187–1202.

Boullier A-M, Ohtani T, Fujimoto K, Ito H, and Dubois M (2001) Fluid inclusions in pseudotachylytes from the Nojima fault, Japan. *Journal of Geophysical Research* 106(B10): 21965–21978. http://dx.doi.org/10.1029/2000JB000043.

Boullier A-M and Robert F (1992) Paleoseismic events recorded in Archaean gold-quartz vein networks, Val d'Or, Abitibi, Quebec, Canada. *Journal of Structural Geology* 14: 161–179.

Brace WF and Kohlstedt DL (1980) Limits on lithospheric stress imposed by laboratory experiments. *Journal of Geophysical Research* 85: 6248–6252.

Brodsky EE and Kanamori H (2001) The elastohydrodynamic lubrication of faults. *Journal of Geophysical Research* 106: 16357–16374.

Brudy M, Zoback MD, Fuchs K, Rummel F, and Baumgärtner J (1997) Estimation of the complete stress tensor to 8 km depth in the KTB scientific drill holes: Implications for crustal strength. *Journal of Geophysical Research* 102: 18453–18475.

Bruhn RL, Parry WT, Yonkee WA, and Thompson T (1994) Fracturing and hydrothermal alteration in normal fault zones. *Pure and Applied Geophysics* 142: 609–644.

Bruhn RL, Yonkee WA, and Parry WT (1990) Structural and fluid-chemical properties of seismogenic normal faults. *Tectonophysics* 175: 139–157.

Brune JN (1970) Tectonic stress and the spectra of seismic shear waves from earthquakes. *Journal of Geophysical Research* 75: 4997–5009.

Brune JN, Brown S, and Johnson PA (1993) Rupture mechanism and interface separation in foam rubber models of earthquakes: A possible solution to the heat flow paradox and the paradox of large overthrusts. *Tectonophysics* 218: 59–67.

Brune JN, Henyey TL, and Roy RF (1969) Heat flow, stress, and rate of slip along the San Andreas Fault, California. *Journal of Geophysical Research* 74: 3821–3827.

Burford RO and Harsh PW (1980) Slip on the San Andreas Fault in central California from alignment array surveys. *Bulletin of the Seismological Society of America* 70: 1233–1261.

Byerlee JD (1978) Friction of rocks. *Pure and Applied Geophysics* 116: 615–629.

Byerlee JD (1990) Friction, overpressure and fault normal compression. *Geophysical Research Letters* 17: 2109–2112.

Byerlee JD (1993) A model for episodic flow of high pressure water in fault zones before earthquakes. *Geology* 21: 303–306.

Byerlee JD and Summers R (1976) A note on the effect of fault gouge thickness on fault stability. *International Journal of Rock Mechanics and Mining Sciences and Geomechanical Abstracts* 13: 35–36.

Byrne T and Fisher D (1990) Evidence for a weak and overpressured décollement beneath sediment-dominated accretionary prisms. *Journal of Geophysical Research* 98: 9081–9097.

Chester FM, Evans JP, and Biegel RL (1993) Internal structure and weakening mechanisms of the San Andreas fault. *Journal of Geophysical Research* 98: 771–786.

Chester FM and Higgs NG (1992) Multimechanism friction constitutive model for ultrafine quartz gouge at hypocentral conditions. *Journal of Geophysical Research* 97: 1859–1870.

Cornet FH, Bernard P, and Moretti I (2004) The Corinth Rift Laboratory. *Comptes Rendus Geosciences* 336: 235–241.

Dahlen FA (1990) Critical taper model of fold-and-thrust belts and accretionary wedges. *Annual Review of Earth and Planetary Sciences* 18: 55–99.

Davis D, Suppe J, and Dahlen FA (1983) The mechanics of fold and thrust belts and accretionary wedges. *Journal of Geophysical Research* 88: 1153–1172.

Dewers T and Ortoleva P (1994) Nonlinear dynamical aspects of deep basin hydrology: Fluid compartment formation and episodic fluid release. *American Journal of Science* 294: 713–755.

Dieterich JH (1992) Earthquake nucleation on faults with rate- and state-dependent strength. *Tectonophysics* 211: 115–134.

Dreger DS, Tkalcic H, and Johnston M (2000) Dilational processes accompanying earthquakes in the Long Valley Caldera. *Science* 288: 122–125.

Eichhubl P and Boles JR (2000) Rates of fluid flow in fault systems – Evidence for episodic rapid fluid flow in the Miocene Monterey Formation Coastal California. *American Journal of Science* 300: 571–600.

Ellsworth WL and Beroza GC (1995) Seismic evidence for an earthquake nucleation phase. *Science* 268: 851–855.

Ellsworth WL and Beroza GC (1998) Observation of the seismic nucleation phase in the Ridgecrest, California earthquake sequence. *Geophysical Research Letters* 25: 401–404.

Evans JP and Chester FM (1995) Fluid–rock interaction in rocks of the San Andreas system: Inferences from San Gabriel fault rock geochemistry and microstructures. *Journal of Geophysical Research* 100: 13007–13020.

Fehler MC (1989) Stress control of seismicity patterns observed during hydraulic fracturing experiments at the Fenton Hill hot dry rock geothermal energy site, New Mexico. *International Journal of Rock Mechanics and Mining Sciences* 26: 211–219.

Fisher DM (1996) Fabrics and veins in the forearc: A record of cyclic fluid flow at depths of <15 km. In: Bebout GE, Schou DW, Kirby SH, and Platt JJ (eds.) *Subduction: Top to Bottom*. *American Geophysical Union Geophysical Monograph*, vol. 96, pp. 75–89. Washington, DC: American Geophysical Union.

Fitzenz DD and Miller SA (2003) Fault compaction and overpressured faults: Results from a 3D model of a ductile fault zone. *Geophysical Journal International* 155: 111–125.

Flesch LM, Holt WE, Haines AJ, and Shen-Tu B (2000) Dynamics of the Pacific-North American plate boundary in the Western United States. *Science* 287: 834–836.

Fournier RL (1996) Compressive and tensile failure at high fluid pressure where preexisting fractures have cohesive strength, with application to the San Andreas fault. *Journal of Geophysical Research* 101: 25499–25509.

Fujimoto K, Tanaka H, Higuchi T, Tomida N, Ohtani T, and Ito H (2001) Alteration and mass transfer inferred from the Hirabayashi GSJ drill penetrating the Nojima fault, Japan. *The Island Arc* 10: 401–410.

Fyfe S, Price NJ, and Thomson AB (1978) *Fluids in the Earth's Crust*. New York: Elsevier.

Hacker BR, Blanpied ML, Lockner DA, and Jove CF (1995) Dehydration and friction: Laumontite → wairakite + H_2O. *Geological Society of America, Abstracts with Programs* 27: A282.

Hainzl S and Ogata Y (2005) Detecting fluid signals in seismicity data through statistical earthquake modeling. *Journal of Geophysical Research* 110: B05S07. http://dx.doi.org/10.1029/2004JB003247.

Harris RA (2004) Numerical simulations of large earthquakes: Dynamic rupture propagation on heterogeneous faults. *Pure and Applied Geophysics* 161: 2171–2181. http://dx.doi.org/10.1007/s00024-004-2556-8.

Healy JH, Rubey WW, Griggs DT, and Raleigh CB (1968) The Denver earthquakes. *Science* 161: 1301–1310.

Heaton TH (1990) Evidence for and implications of self-healing pulses of slip in earthquake rupture. *Physics of the Earth and Planetary Interiors* 64: 1–20.

Heesakkers V, Murphy SK, van Aswegen G, et al. (2005) The rupture zone of the $M = 2.2$ earthquake that reactivated the ancient Pretorius Fault in TauTona Mine, South Africa. *EOS. Transactions of the American Geophysical Union* 86, S31B-04.

Henyey TL and Wasserburg GJ (1971) Heat flow near major strike-slip faults in California. *Journal of Geophysical Research* 76: 7924–7946.

Hickman S (1991) Stress in the lithosphere and the strength of active faults, US National Report on International Union of Geodesy and Geophysics 1987–1990. *Reviews of Geophysics* 29: 759–775.

Hickman S, Barton CA, Zoback MD, Morin R, Sass J, and Benoit R (1997) *In-situ* stress and fracture permeability along the Stillwater fault zone, Dixie Valley, Nevada. *International Journal of Rock Mechanics and Mining Sciences* 34: 3–4 (Paper No. 126).

Hickman S and Evans B (1992) Growth of grain contacts in halite by solution transfer: Implications for diagenesis, lithification, and strength recovery. In: Evans B and Wong TF (eds.) *Fault Mechanics and Transport Properties of Rocks*, pp. 253–280. San Diego, CA: Academic Press.

Hickman S, Sibson R, and Bruhn R (1995) Introduction to special section: Mechanical involvement of fluids in faulting. *Journal of Geophysical Research* 100: 12831–12840.

Hickman S and Zoback MD (2004) Stress measurements in the SAFOD pilot hole: Implications for the frictional strength of the San Andreas fault. *Geophysical Research Letters* 31: L15S12.

Hickman S, Zoback MD, and Ellsworth WE (2004) Introduction to special section: Preparing for the San Andreas Fault Observatory at Depth. *Geophysical Research Letters* 31: L12S01. http://dx.doi.org/10.1029/20040GL20688.

Hubbert MK and Rubey WW (1959) Role of fluid pressure in mechanics of overthrust faulting. *Geological Society of America Bulletin* 70: 115–205.

Hunt JM (1990) Generation and migration of petroleum from abnormally pressured fluid compartments. *American Association of Petroleum Geologists Bulletin* 74: 1–12.

Iio Y (1992) Slow initial phase of the S-wave velocity pulse generated by microearthquakes. *Geophysical Research Letters* 19: 477–480.

Iio Y (1995) Observation of the slow initial phase of microearthquakes implications for earthquake nucleation and propagation. *Journal of Geophysical Research* 100: 15333–15349.

Imanishi K and Ellsworth WL (2006) Source scaling relationships of microearthquakes at Parkfield, CA, determined using the SAFOD pilot hole seismic array. In: McGarr A, Kanamori H, and Abercrombie R (eds.) *AGU Monograph on Radiated Energy and the Physics of Earthquake Faulting. Geophysical Monograph Series*, vol. 170, pp. 81–90. American Geophysical Union.

Imanishi K, Ellsworth WE, and Prejean SG (2004) Earthquake source parameters determined by the SAFOD pilot hole array. *Geophysical Research Letters* 31: L12S09. http://dx.doi.org/10.1029/2004GL019420.

Janecke SU and Evans JP (1988) Feldspar-influenced rock rheologies. *Geology* 16: 1064–1067.

Kanamori H (1980) State of stress in the Earth's lithosphere. In: Dziewonski AM and Boschi E (eds.) *Physics of the Earth's Interior*, pp. 531–554. Bologna, Italy: Societa Italiana di Fisica.

Kano H, Mori J, Fujio R, et al. (2006) Heat signature of the Chelungpu fault associated with the 1999 Chi-Chi Taiwan earthquake. *Geophysical Research Letters* 33, L14306.

Kao H and Angelier J (2001) Stress tensor inversion for the Chi-Chi earthquake sequence and its implications on regional collision. *Bulletin of the Seismological Society of America* 91: 1028–1040.

Karner S, Marone C, and Evans B (1997) Laboratory study of fault healing and lithification in simulated fault gouge under hydrothermal conditions. *Tectonophysics* 277: 41–55.

Kennedy BM, Kharaka YH, Evans WC, et al. (1997) Mantle fluids in the San Andreas fault system, California. *Science* 278: 1278–1281.

Kerrich R, La Tour TE, and Willmore L (1984) Fluid participation in deep fault zones: Evidence from geological, geochemical, and $^{18}O/^{16}O$ relations. *Journal of Geophysical Research* 89: 4331–4343.

Kirby SH (1983) Rheology of the lithosphere. *Reviews of Geophysics and Space Physics* 21: 1458–1487.

Ko S-Z, Olgaard DL, and Wong T-F (1997) Generation and maintenance of pore pressure excess in a dehydrating system 1. Experimental and microstructural observations. *Journal of Geophysical Research* 102: 825–839.

Kostrov BV (1964) Self-similar problem of propagation of shear cracks. *Journal of Applied Mathematics and Mechanics* 28: 1077–1087.

Lachenbruch AH (1980) Frictional heating, fluid pressure, and the resistance to fault motion. *Journal of Geophysical Research* 85: 6097–6112.

Lachenbruch AH and McGarr A (1990) Stress and heat flow. In: Wallace RE (ed.) *US Geological Survey Professional Paper 1515: The San Andreas Fault System, California*, pp. 261–277. Menlo Park, CA: USGS.

Lachenbruch AH and Sass JH (1980) Heat flow and energetics of the San Andreas fault zone. *Journal of Geophysical Research* 85: 6185–6223.

Lachenbruch AH and Sass JH (1992) Heat flow from Cajon Pass, fault strength and tectonic implications. *Journal of Geophysical Research* 97: 4995–5015.

Lachenbruch AH and Thompson GA (1972) Oceanic ridges and transform faults: Their intersection angles and resistance to place motion. *Earth and Planetary Science Letters* 15: 116–122.

Lockner DA and Byerlee JD (1995) An earthquake instability model based on faults containing high fluid-pressure compartments. *Pure and Applied Geophysics* 145: 717–745.

Logan JM and Rauenzahn KA (1987) Frictional dependence of gouge mixtures of quartz and montmorillonite on velocity, composition and fabric. *Tectonophysics* 144: 87–108.

Lund B and Zoback MD (1999) State of stress to 7 km depth in the Siljan drill holes from observations of drilling-induced tensile wall fractures. *International Journal of Rock Mechanics and Mining Sciences* 36: 169–190.

Ma K-F, Brodsky EE, Mori J, Song C, Ji TR, and Kanamori H (2003) Evidence for fault lubrication during the 1999 Chi-Chi, Taiwan, earthquake (Mw 7.6). *Geophysical Research Letters* 30: 1244.

Ma K-F, Tanaka H, Song S-R, et al. (2006) Slip zone and energetics of a large earthquake from the Taiwan Chelungpu-fault Drilling Project. *Nature* 444: 473–476.

Madariaga R (1976) Dynamics of an expanding circular fault. *Bulletin of the Seismological Society of America* 66: 639–666.

Magee ME and Zoback MD (1993) Evidence for a weak interpolate thrust fault along the northern Japan subduction zone and implications for the mechanics of thrust faulting and fluid expulsion. *Geology* 21: 809–812.

Marone C, Raleigh CB, and Scholz CH (1990) Frictional behavior and constitutive modeling of simulated fault gouge. *Journal of Geophysical Research* 95: 7007–7025.

Martinsen RS (1997) Stratigraphic controls on the development and distribution of fluid-pressure compartments. *American Association of Petroleum Geologists Memoir* 67: 223–241.

Mase CW and Smith L (1987) Effects of frictional heating on the thermal, hydrological, and mechanical response of a fault. *Journal of Geophysical Research* 92: 6249–6272.

McNutt M (1980) Implications of regional gravity for state of stress in the Earth's crust and upper mantle. *Journal of Geophysical Research* 85: 6377–6396.

McNutt MK and Menard HW (1982) Constraints on yield strength in the oceanic lithosphere derived from observations of flexure. *Geophysical Journal of the Royal Astronomical Society* 71: 363–394.

Melosh HJ (1979) Acoustic fluidization: A new geologic process? *Journal of Geophysical Research* 84: 7513–7520.

Melosh HJ (1996) Dynamic weakening of faults by acoustic fluidization. *Nature* 379: 601–606.

Miller SA (1996) Fluid-mediated influence of adjacent thrusting on the seismic cycle at Parkfield. *Nature* 382: 799–802.

Moore DE, Lockner DA, Summers R, Shengli M, and Byerlee JD (1997) Strengths of serpentinite gouges at elevated temperatures. *Journal of Geophysical Research* 102: 14787–14801.

Moore JC, Shipley TH, Goldberg D, et al. (1995) Abnormal fluid pressure and fault-zone dilation in the Barbados accretionary prism: Evidence from logging while drilling. *Geology* 23: 605–608.

Moore JC and Vrolijk P (1992) Fluids in accretionary prisms. *Reviews of Geophysics* 30: 113–135.

Morrow C, Radney B, and Byerlee J (1992) Frictional strength and the effective pressure law of montmorillonite and illite clays. In: Evans B and Wong T-F (eds.) *Fault Mechanics and Transport Properties of Rocks*, pp. 69–88. San Diego, CA: Academic Press.

Morrow CA, Shi LQ, and Byerlee JD (1982) Strain hardening and strength of clay-rich fault gouges. *Journal of Geophysical Research* 87: 6771–6780.

Mount VS and Suppe J (1987) State of stress near the San Andreas Fault: Implications for wrench tectonics. *Geology* 15: 1143–1146.

Mount VS and Suppe J (1992) Present-day stress orientations adjacent to active strike-slip faults: California and Sumatra. *Journal of Geophysical Research* 97: 11995–12013.

Mrotek KA, Quittmeyer RC, Naumhoff PG, and Stratton CT (1988) Observations of the earthquake swarm near Moodus, Connecticut. September/October, 1987. *EOS. Transactions of the American Geophysical Union* 69: 495.

Nadeau R, Antolik M, Johnson P, Foxall W, and McEvilly TV (1994) Seismological studies at Parkfield III: Microearthquake clusters in the study of fault-zone dynamics. *Bulletin of the Seismological Society of America* 84: 247–263.

Nadeau RM, Foxall W, and McEvilly TV (1995) Clustering and periodic recurrence of microearthquakes on the San Andreas fault at Parkfield, California. *Science* 267: 503–507.

Nadeau RM and McEvilly TV (1997) Seismological studies at Parkfield V: Characteristic microearthquake sequences as fault-zone drilling targets. *Bulletin of the Seismological Society of America* 87: 1463–1472.

Nadeau RM, Michelini A, Uhrhammer RA, Dolenc D, and McEvilly TV (2004) Detailed kinematics, structure and recurrence of micro-seismicity in the SAFOD target region. *Geophysical Research Letters* 31: L12S08. http://dx.doi.org/10.1029/2003GL019409.

Nur A and Booker JR (1972) Aftershocks caused by pore-fluid flow? *Science* 175: 885–887.

Ohnaka M (1992) Earthquake source nucleation: A physical model for short-term precursors. *Tectonophysics* 211: 149–178.

Ohtani T, Tanaka H, Fujimoto K, Higuchi T, Tomida N, and Ito H (2001) Internal structure of the Nojima fault zone from the Hirabayashi GSJ drill core. *The Island Arc* 10: 392–400.

Oppenheimer DH, Reasenberg PA, and Simpson RW (1988) Fault-plane solutions for the 1984 Morgan Hill, California earthquake sequence: Evidence for the state of stress on the Calaveras fault. *Journal of Geophysical Research* 93: 9007–9026.

Oshiman N, Shimamoto T, and Takemura K, et al. (eds.) (2001) Nojima Fault Zone probe. *The Island Arc* 10: 195–505.

Parry WT (1994) Fault fluid compositions from fluid inclusion observations. In: Hickman S, Sibson R, and Bruhn R (eds.) *Proceedings USGS Red Book Conference on the Mechanical Involvement of Fluids in Faulting*, pp. 334–348. Menlo Park, CA: USGS.

Peyrat S, Olsen KB, and Madariaga R (2004) Which dynamic rupture parameters can be estimated from strong ground motion and geodetic data? *Pure and Applied Geophysics* 161: 2155–2169. http://dx.doi.org/10.1007/s00024-004-2555-9.

Pinkston J, Stern L, and Kirby S (1987) Hydrothermal reactions on artificial fault surfaces in dunite: Fibrous mineral growth, slickensides and temperature sensitivity of reaction weakening. *EOS. Transactions of the American Geophysical Union* 68: 405.

Power WL and Tullis TE (1989) The relationship between slickenside surfaces in fine-grained quartz and the seismic cycle. *Journal of Structural Geology* 11: 879–893.

Powley DE (1990) Pressures and hydrogeology in petroleum basins. *Earth Science Reviews* 29: 215–226.

Prejean SG and Ellsworth WL (2001) Observations of earthquake source parameters at 2 km depth in the Long Valley Caldera, Eastern California. *EOS. Transactions of the American Geophysical Union* 91: 165–177.

Raleigh CB, Healy JH, and Bredehoeft JD (1972) An experiment in earthquake control at Rangely, Colorado. *Science* 191: 1230–1237.

Reches Z, DAFSAM, and NELSAM Teams (2006) Building a natural earthquake laboratory at focal depth (DAFSAM–NELSAM Project, South Africa). *Scientific Drilling* 3: 30–33.

Reinen LA and Tullis TE (1995) Microstructural evidence of strain localization and distributed strain in serpentine friction experiments. *EOS. Transactions of the American Geophysical Union* 76: F560.

Reinen LA, Weeks JD, and Tullis TE (1994) The frictional behavior of lizardite and antigorite serpentinites: Experiments, constitutive models, and implications for natural faults. *Pure and Applied Geophysics* 143: 317–358.

Rice JR (1992) Fault stress states, pore pressure distributions, and the weakness of the San Andreas fault. In: Evans B and Wong T-F (eds.) *Fault Mechanics and Transport Properties of Rocks*, pp. 475–503. San Diego, CA: Academic Press.

Rice JR (2006) Heating and weakening of faults during earthquake slip. *Journal of Geophysical Research* 111: B05311. http://dx.doi.org/10.1029/2005JB004006, 29 pages.

Rice JR and Cleary MP (1976) Some basic stress diffusion solutions for fluid-saturated elastic porous media with compressible constituents. *Reviews of Geophysics* 14: 227–241.

Rudnicki J (1988) Physical models of earthquake instability and precursory processes. *Pure and Applied Geophysics* 126: 524–531.

Rutter EH and Mainprice DH (1979) On the possibility of slow fault slip controlled by diffusive mass transfer processes. *Gerlands Beitrage zur Geophysik* 88: 154–162.

Saffer DM and Bekins BA (2006) An evaluation of factors influencing pore pressure in accretionary complexes: Implications for taper angle and wedge mechanics. *Journal of Geophysical Research* 111: B04101. http://dx.doi.org/10.1029/2005JB003990.

Saffer DM, Bekins BA, and Hickman S (2003) Topographically driven groundwater flow and the San Andreas heat flow paradox revisited. *Journal of Geophysical Research* 108(B5): 2274.

Scholz CH (1989) Mechanics of faulting. *Annual Review of Earth and Planetary Sciences* 17: 309–334.

Scholz CH (2000) Evidence for a strong San Andreas fault. *Geology* 28(2): 163–166.

Schulz SS (1989) Catalog of creepmeter measurements in California from 1966 through 1988. *US Geological Survey Open-File Report 89-650*. Menlo Park, CA: USGC.

Schulz SE and Evans JP (2000) Mesoscopic structure of the Punchbowl fault, southern California and the geologic and geophysical structure of active strike-slip faults. *Journal of Structural Geology* 22: 913–930.

Segall P and Rice JR (1995) Dilatancy, compaction and slip instability of a fluid-infiltrated fault. *Journal of Geophysical Research* 100: 22155–22171.

Shimamoto T and Logan J (1981) Effects of simulated clay gouges on the sliding behavior of Tennessee sandstone. *Tectonophysics* 75: 243–255.

Shipley TH, Moore GF, Bangs NL, Moore JC, and Stoffa PL (1994) Seismically inferred dilatancy distribution, northern Barbados Ridge decollement: Implications for fluid migration and fault strength. *Geology* 22: 411–414.

Sibson RH (1973) Interactions between temperature and pore-fluid pressure during earthquake faulting: A mechanism for partial or total stress relief. *Nature Physical Science* 243: 66–68.

Sibson RH (1985) Stopping of earthquake ruptures at dilational fault jogs. *Nature* 316: 248–251.

Sibson RH (1990) Rupture nucleation on unfavorably oriented faults. *Bulletin of the Seismological Society of America* 80: 1580–1604.

Sleep NH (1995) Ductile creep, compaction and rate and state dependent friction within major fault zones. *Journal of Geophysical Research* 100: 13065–13080.

Sleep NH and Blanpied ML (1992) Creep, compaction, and the weak rheology of major faults. *Nature* 359: 687–692.

Sleep NH and Blanpied ML (1994) Ductile creep and compaction: A mechanism for transiently increasing fluid pressure in mostly sealed fault zones. *Pure and Applied Geophysics* 143: 9–40.

Smith L, Forster CB, and Evans JP (1990) Interaction of fault zones, fluid flow and heat transfer on a basin scale. In: Newman SP and Neretnieks I (eds.) *Hydrogeology of Low Permeability Environments*, pp. 41–67. Wallingford, UK: International Association of Hydrological Sciences, Selected Papers, 2.

Spray JG (1987) Artificial generation of pseudotachylite using friction welding apparatus: Simulation of melting on a fault plane. *Journal of Structural Geology* 9: 49–60.

Stock JM, Healy JH, Hickman SH, and Zoback MD (1985) Hydraulic fracturing stress measurements at Yucca Mountain, Nevada, and relationship to regional stress field. *Journal of Geophysical Research* 90: 8691–8706.

Tanaka H, Chen WM, Kawabata K, and Urata N (2007) Thermal properties across the Chelungpu fault zone and evaluations of positive thermal anomaly on the slip zones: Are these residuals of heat from faulting. *Geophysical Research Letters* 34: L1309.

Tanaka H, Kujimoto K, Ohtani T, and Ito H (2001) Structural and chemical characterization of shear zones in the freshly activated Nojima fault, Awaji Island, southwest Japan. *Journal of Geophysical Research* 106: 8789–8810.

Tanaka H, Wang CY, Chen WM, et al. (2002) Initial science report of shallow drilling penetrating into the Chelungpu fault zone, Taiwan. *Terrestrial, Atmospheric and Oceanic Sciences* 13: 227–251.

Townend J and Zoback MD (2000) How faulting keeps the crust strong. *Geology* 28(5): 399–402.

Townend J and Zoback MD (2001) Implications of earthquake focal mechanisms for the frictional strength of the San Andreas fault system. In: Holdsworth RE, Strachan RA, Macloughlin JJ, and Knipe RJ (eds.) *The Nature and Tectonic Significance of Fault Zone Weakening. Special Publication of the Geological Society of London*, vol. 186, pp. 13–21. London: Geological Society of London.

Townend J and Zoback MD (2004) Regional tectonic stress near the San Andreas fault in Central and Northern California. *Geophysical Research Letters* 31(13): L15S11.

Wakita H and Sano Y (1987) High ^{3}He emanation and seismic swarms observed in a nonvolcanic, forearc region. *Journal of Geophysical Research* 92: 12539–12546.

Wald DJ and Heaton TH (1994) Spatial and temporal distribution of slip for the 1992 Landers, California, earthquake. *Bulletin of the Seismological Society of America* 84: 668–691.

Wang CY, Mao N, and Wu FT (1980) Mechanical properties of clays at high pressure. *Journal of Geophysical Research* 85: 1462–1468.

Wang K, Mulder T, Rogers G, and Hyndman R (1995) Case for very low coupling stress on the Cascadia subduction fault. *Journal of Geophysical Research* 100: 12907–12918.

Wilcock WSD, Purdy GM, and Solomon SC (1990) Microearthquake evidence for extension across the Kane transform fault. *Journal of Geophysical Research* 95: 15439–15462.

Williams CF, Grubb FV, and Galanis SP (2004) Heat flow in the SAFOD pilot hole and implications for the strength of the San Andreas Fault. *Geophysical Research Letters* 31: L15S14. http://dx.doi.org/10.1029/2003GL019352.

Wintsch R, Christoffersen R, and Kronenberg A (1995) Fluid–rock reaction weakening of fault zones. *Journal of Geophysical Research* 100: 13021–13032.

Wu FT (1978) Mineralogy and physical nature of clay gouge. *Pure and Applied Geophysics* 116: 655–689.

Wu H-Y, Ma K-F, Zoback M, Boness N, Ito H, and Hung J-H (2007) Stress orientations of Taiwan Chelungpu-Fault Drilling Project (TCDP) hole-A as observed from geophysical logs. *Geophysical Research Letters* 34, L01303.

Zoback MD (2000) Strength of the San Andreas. *Nature* 405: 31–32.

Zoback MD, Barton CA, Brudy M, et al. (2003) Determination of stress orientation and magnitude in deep wells. *International Journal of Rock Mechanics and Mining Sciences* 40: 1049–1076.

Zoback MD and Beroza GC (1993) Evidence for near-frictionless faulting in the 1989 (*M* 6.9) Loma Prieta, California, earthquake and its aftershocks. *Geology* 21: 181–185.

Zoback MD and Healy JH (1984) Friction, faulting, and '*in situ*' stresses. *Annales Geophysicae* 2: 689–698.

Zoback MD, Zoback ML, Mount VS, et al. (1987) New evidence on the state of stress of the San Andreas fault system. *Science* 238: 1105–1111.

This page intentionally left blank

4.08 Dynamic Shear Rupture in Frictional Interfaces: Speeds, Directionality, and Modes

AJ Rosakis, California Institute of Technology, Pasadena, CA, USA
K Xia, University of Toronto, Toronto, ON, Canada
G Lykotrafitis, Massachusetts Institute of Technology, Cambridge, MA, USA
H Kanamori, California Institute of Technology, Pasadena, CA, USA

© 2007 Elsevier B.V. All rights reserved.

This chapter is reproduced from the previous edition, Volume 4, pp. 153–192, © 2007, Elsevier B.V.

4.08.1	**Introduction**	183
4.08.2	**Experimental Techniques for Creating Earthquakes in the Laboratory**	184
4.08.2.1	Early Experimental Studies	184
4.08.2.2	Experimental Design Philosophy	185
4.08.2.2.1	Simulating tectonic loading	185
4.08.2.2.2	Simulating the nucleation process	186
4.08.2.2.3	Choosing appropriate model materials and diagnostics	187
4.08.3	**Supershear and Sub-Rayleigh to Supershear Transition in Homogeneous Fault Systems**	187
4.08.3.1	Purely Sub-Rayleigh and Purely Supershear Earthquake Ruptures	189
4.08.3.2	The Experimental Visualization of the Sub-Rayleigh to Supershear Earthquake Rupture Transition	191
4.08.3.3	A Theoretical Model for the Sub-Rayleigh to Supershear Transition	192
4.08.3.4	Discussion	193
4.08.4	**Directionality of Ruptures Along Faults Separating Weakly Dissimilar Materials: Supershear and Generalized Rayleigh Wave Speed Ruptures**	194
4.08.4.1	Two Types of Ruptures Along Inhomogeneous Fault Systems	194
4.08.4.2	Experimental Setup	196
4.08.4.3	Experimental Results	196
4.08.4.4	Comparison of the Experimental Results to Early Numerical and Theoretical Studies	198
4.08.4.5	The Parkfield Earthquake Discussion in the Context of Experiments and of Recent Numerical Studies	199
4.08.4.6	Discussion of the Historic North Anatolian Fault Earthquake Sequence in View of the Experimental Results	200
4.08.5	**Observing Crack-Like, Pulse-Like, Wrinkle-Like, and Mixed Rupture Modes in the Laboratory**	201
4.08.5.1	Experimental Investigation of Dynamic Rupture Modes	202
4.08.5.1.1	Specimen configuration and loading	203
4.08.5.1.2	Using particle velocimetry to measure slip and opening histories	203
4.08.5.2	Visualizing Pulse-Like and Crack-Like Ruptures in the Laboratory	204
4.08.5.3	Wrinkle-Like Opening Pulses Along Interfaces in Homogeneous Systems	208
4.08.5.4	Discussion	210
Acknowledgments		210
References		210

4.08.1 Introduction

The dramatic progress of research on earthquake rupture in the last decades along with advances in the various branches of geophysics has decisively improved our understanding of seismic source events. A tremendous scientific activity including theoretical seismomechanics studies, field data processing, and laboratory investigations has answered many outstanding questions on earthquake dynamics. Other problems yet remain to be solved whereas new questions have recently emerged. One of these questions is related to the ability of laboratory experiments to faithfully 'mimic' natural earthquake rupture events. In particular, this question is concerned with the 'scalability' of the laboratory results and observations. Can conclusions regarding the nature of dynamic rupture (e.g., modes, speeds, directionality, etc.), obtained from such experiments, be accurately transferred from the laboratory scale (meters) to the natural earthquake rupture scale (hundreds of kilometers)? Another question relates to similar issues in modeling. Do mechanics models and numerical simulations of rupture also suffer from scaling issues and do they need to first be validated by direct comparison to experiments before their results are extrapolated to the scale of rupture processes in the Earth's crust? These are some of the questions which, although only partially addressed here, have motivated much of the work reviewed in this chapter.

In this chapter, we first review past work on experimental simulations of earthquake rupture in the laboratory. We describe the basic traditional experimental techniques used in the last few decades and we then focus on a new experimental design philosophy introduced by the authors. This new approach to laboratory earthquakes has resulted in an experimental configuration particularly suited to the study of fast rupture events.

Following the description of the design philosophy, the chapter also concentrates on experimental results which demonstrate the attainability of intersonic rupture speeds (i.e., speeds in the interval between the shear and the dilatational wave speeds) and elucidate the sub-Rayleigh to supershear transition phenomenon in homogeneous fault systems. This transition is discussed in the light of well-established theoretical and numerical models based on the principles of fracture mechanics. In addition, recent field evidence of intersonic rupture that has occurred during natural earthquakes is presented and compared to the experimental observations.

The issue of the directionality of ruptures along faults separating weakly dissimilar materials is currently in the center of a very active debate between seismologists. In this chapter, theoretical analyses together with field evidence and recent experimental and numerical results are utilized to provide a complete picture of the state of contemporary knowledge on this problem.

Finally, the various rupture modes (crack-like, pulse-like, wrinkle-like) which have been proposed on the basis of field evidence and theoretical analyses are critically discussed. Experimental results, which show that the activation of these modes in the laboratory scale is possible, are presented.

Here, we have made the conscious decision not to include an extensive general introduction to this article. Instead, we have chosen to strengthen the introductory comments of individual sections and to blend reference to previous and current work within each section.

4.08.2 Experimental Techniques for Creating Earthquakes in the Laboratory

4.08.2.1 Early Experimental Studies

The first efforts of simulating earthquakes in the laboratory have utilized the classic block-slider configuration. Based on valuable observations gathered through this configuration, Brace and Byerlee (1966) have proposed that stick–slip may be an important mechanism for shallow earthquakes occurring along preexisting faults. The basic assumption of the block-slider model is that the sliding rock blocks are rigid so that edge effects, leading to an initially nonuniform stress distribution along the interface, can be ignored. Under the rigidity assumption, friction along the interface can be assumed uniform and as a result only the displacement history of the slider needs to be measured. This assumption is clearly a very strong one since the Earth's crust is far from being rigid during earthquake faulting. Although the block-slider model experiments do provide useful insight into some aspects of earthquake mechanics, such models are more appropriate for providing frictional constitutive information instead of modeling the detailed mechanics of earthquake rupture. Indeed, based on the pioneering experimental results of Dieterich (1979), the famous state- and rate-dependent friction law was formulated by Ruina (1983) and Rice (1983).

There are two configurations suitable for producing laboratory earthquakes, namely those of direct shear and of biaxial compression (Brune, 1973; Johnson and Scholz, 1976; Wu et al., 1972). In order to extract meaningful rupture dynamics information from these configurations, the assumption of rigidity should be abandoned. Indeed, if strain gauges are attached along the fault in a two-dimensional (2-D) domain, the dynamic strain history resulting from the rupture can be recorded at various points. With this modification, rupture velocities can also be estimated (Brune, 1973; Johnson and Scholz, 1976; Wu et al., 1972) in addition to the dynamic friction-slip relation (Dieterich and Kilgore, 1994; Ohnaka and Shen, 1999; Okubo and Dieterich, 1984). However, strain gauges have approximate sizes of few millimeters, and this limits the spatial resolution of these discrete measurements. As a result, the data are pointwise, averaged, and are arguably difficult to interpret. For example, although supershear ruptures were suggested by this type of experiment (Johnson and Scholz, 1976; Okubo and Dieterich, 1984; Wu et al., 1972), this evidence was considered far from being conclusive due to the 'averaged' quality of the data. Furthermore, since the diagnostics system (the Wheatstone bridge and oscilloscope) was triggered by one of the strain gauge signals and there was no information on the exact time and location of the nucleating point of earthquake (hypocenter), the sampling rate was necessarily set too low to capture the whole process. This has limited the time resolution of the measurements. Despite their shortcomings, these early studies have produced very valuable information concerning the nature and modes of frictional sliding. Indeed, by using this type of setup, the two important phenomena of frictional healing and velocity weakening were first observed by Scholz et al. (1972) and Dieterich (1972).

Recently, two important modifications of the original block-slider model were introduced. The first one involved introducing a granular layer into the sliding interface to simulate the fault core (Gu and Wong, 1994). The second one involved using soft materials (e.g., foam rubber) and multipoint measurements performed by burying accelerometers inside meter-sized blocks (Brune, 1973; Brune et al., 1993). The first type of test usually ignored the dynamic features of the earthquake, and is thus beyond the scope of this chapter. As to the second type of test, namely, the 'foam model,' the usage of slow wave speed material has some distinct advantages over using rocks in that: (1) the whole process can be captured without the need of high sampling rates or high-speed photography and (2) 3-D effects can be more effectively addressed, since it is relatively easy to bury gauges inside the model. However, some concerns exist regarding the model material itself. These include its nonlinear and viscous nature and the special frictional properties of the foam rubber (the coefficient of friction can be larger than unity). Regarding their constitutive behavior, foam rubbers can sustain large deformations and are not well described by the theory of linear elasticity. The presence of material nonlinearities (both kinematic and constitutive) raises concerns regarding the scalability of the results from the laboratory scale to naturally occurring earthquake ruptures. This together with their known viscoelastic response makes foam models unlikely candidates to simulate earthquakes, which occur in much stiffer and more brittle materials (i.e., rocks) and also involve sliding processes extending over lengths that are four to five orders of magnitude longer than in the laboratory. Furthermore, like all other direct shear-type experiments, the 'foam model' has also suffered from unavoidable edge effects as first pointed out by Scholz et al. (1972).

Despite the few shortcomings described above, it should be emphasized that the foam rubber model has produced very interesting observations whose scientific value has been to elucidate some of the basic physics governing the sliding of frictional interfaces (Anooshehpoor and Brune, 1994, 1999; Brune and Anooshehpoor, 1997; Brune et al., 1993).

At least three important simultaneously acting prerequisites are needed in order for an earthquake to occur either in the laboratory or in the Earth's crust. These are proper tectonic loading, a preexisting fault, and the occurrence of a certain type of triggering mechanism. The first two prerequisites were very well simulated and controlled in the previously mentioned block slider and foam rubber models. However, the triggering of the rupture was mainly left uncontrolled and was dominated by the configuration edge effects discussed above. The triggering mechanism is important for both the earthquake process and its subsequent evolution (Lu et al., 2005). In experiments, control of the triggering mechanism is crucial since it effects the synchronization of high-speed diagnostics and also serves as an important boundary/initial conditions that may affect the repeatability and the interpretation of the results. Except for one particular study specifically targeting the problem of the dynamic triggering of earthquakes by incoming Rayleigh waves (Uenishi et al., 1999), this aspect has largely been ignored by the previous literature. In this important investigation, two photoelastic polymer plates of different sizes were held together at their convex sides to form a fault (line of frictional contact) with an extended free surface. A Rayleigh wave pulse was subsequently generated on the free surface by a point explosion and this wave propagated along the free surface toward the contact part (simulated fault). The incoming Rayleigh pulse was used to trigger dynamic slip along the fault in a controlled and repeatable fashion. In addition to controlling rupture initiation, these experiments were also the first ones to feature the use of full-field optical diagnostics to study the dynamic rupture process in 'incoherent' (frictional) interfaces.

In addition to the experimental studies described above, there also exists a class of dynamic shear fracture experiments originating from the engineering community which are closely related to the phenomenon of geological fault rupture. As reviewed by Rosakis (2002), fracture mechanicians are recently paying more and more attention to the dynamic shear rupture propagating along material interfaces which are becoming common in modern engineering structures and composite materials. These interfaces usually feature finite strength and toughness and can resist both opening and shear loading. They are often referred to as 'coherent' interfaces. Since they are in general weaker than the constituent matrix material, these interfaces often trap dynamic ruptures, of both the opening and shear types and serve as sites of catastrophic failure. On a much larger length scale, crustal faults provide natural weak frictional interfaces where shear-dominated earthquake ruptures occur under the superimposed action of tectonic pressure. Experimental methods for dynamic fracture mechanics, such as photoelasticity and coherent gradient sensing combined with high-speed photography, have been applied successfully in the experimental study of shear ruptures propagating dynamically along coherent interfaces in engineering composites of various types (Coker and Rosakis, 2001; Coker et al., 2003; Lambros and Rosakis, 1995; Rosakis et al., 1998, 1999, 2000). These two diagnostic techniques, which are 2-D in nature, are able to provide very valuable full-field and real-time stress information. Specifically, the photoelastic method, which measures maximum in-plane shear stress contours, is an attractive candidate for the study of shear-dominated processes such as earthquake ruptures.

Further discussion of experimental methodologies and results pertaining to the investigation of various modes of rupture can be found in **Section 4.08.5.1**. In the following section, we present the principles that have guided the design of simple experiments involving incoherent (frictional) interfaces whose purpose has been to simulate earthquake rupture in a highly controlled laboratory setting.

4.08.2.2 Experimental Design Philosophy

One goal in designing dynamic frictional experiments simulating earthquake rupture has been to create a testing environment or platform which could reproduce some of the basic physics governing the rupture dynamics of crustal earthquakes while preserving enough simplicity so that clear conclusions can be obtained by pure observation. Another goal has been to ensure that this platform would be versatile enough to accommodate increasing levels of modeling complexity. We have concentrated in creating experiments that are well instrumented and feature diagnostics of high temporal and spatial resolution. In order to keep the design as simple as possible, we have decided to first concentrate on experiments involving essentially 2-D processes which may be thought of as 'experimental' equivalents of the multiplicity of 2-D numerical models of rupture which exist in the open literature. Because of its predominant two dimensionality, this configuration is perhaps better suited to mimic processes corresponding to larger earthquake events which have already saturated the seismogenic zone and as such, involve primarily 2-D, mode-II rupture growth. The design of the generic experimental configuration, the earthquake triggering mechanism, and the choice of model materials and diagnostics are guided by the following principles.

4.08.2.2.1 Simulating tectonic loading

As shown in **Figure 1**, the crust is simulated by a 'large' photoelastic plate with a thickness of 9.5 mm and 150 mm × 150 mm in 2-D plane dimensions. The plate is large enough so that during the time window of observation there are no reflected wave arrivals. The plate is cut into two identical quadrilaterals, which are then put together, and the frictional interface is used to simulate a fault. The angle of the fault line to the horizontal is denoted by α while the uniaxial pressure acting at the top and the bottom ends of the sample is denoted by P. Both P and α are system variables. Following these definitions, the resolved shear traction τ and the resolved normal traction σ along the fault can be expressed in terms of angle α and pressure P as

$$\tau = P \sin\alpha \cos\alpha \quad \text{and} \quad \sigma = P \cos^2\alpha \qquad [1]$$

To make the connection with commonly used geophysics terminology, a dimensionless factor s, which is often used by seismologists to describe the loading along faults with respect to the strength of the fault (Scholz, 2002), is introduced. By

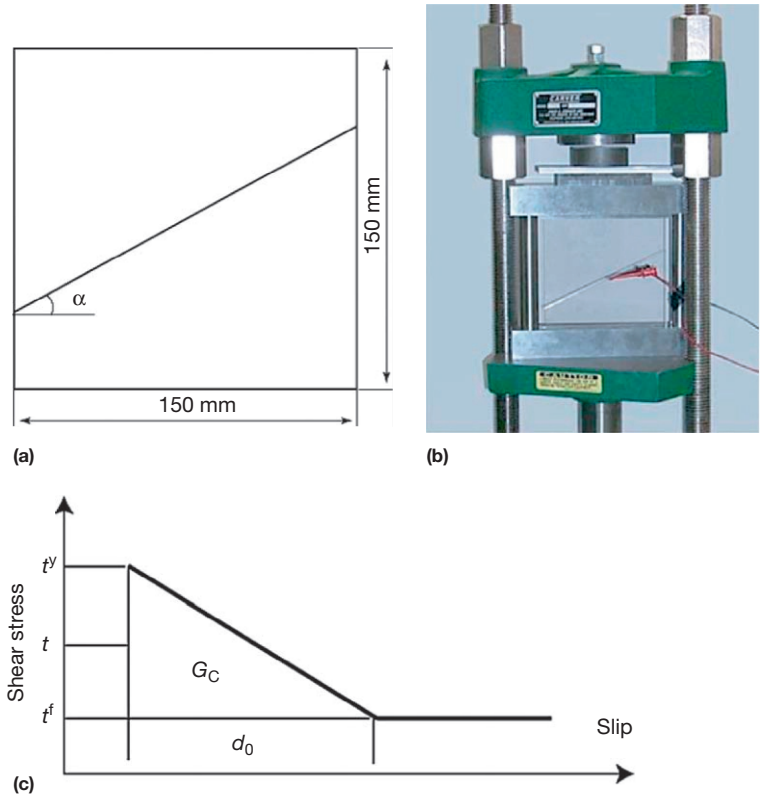


Figure 1 Laboratory fault model (a) and the loading fixture inside a hydraulic press (b). Slip-weakening frictional law (c).

invoking the slip-weakening frictional model (Palmer and Rice, 1973) which is shown schematically in **Figure 1(c)** and by denoting the maximum strength of the fault by τ^y and the final strength of the fault by τ^f, this loading factor (or dimensionless driving force) s is defined by $s=(\tau^y-\tau)/(\tau-\tau^f)$. According to the linear slip-weakening fault strength model, τ^y is the static frictional strength while τ^f is the dynamic frictional strength of the interface. By denoting the static and dynamic coefficients of frictions by μ^s and μ^d, respectively, and by using eqn [1], the loading factor s corresponding to the geometry of the experiment of **Figure 1** becomes

$$s=(\mu^s\cos\alpha-\sin\alpha)/(\sin\alpha-\mu^d\cos\alpha) \qquad [2]$$

As is obvious from eqn [2], s, and through it the earthquake rupturing process, can be controlled by varying the fault angle α for given frictional properties of the fault in the experiments. The magnitude of the uniaxial pressure P does not appear in s. However, it controls the total amount of deformation and the total slip.

4.08.2.2.2 Simulating the nucleation process

Researchers have tried to understand the nucleation process of earthquakes (triggering earthquake rupture) and indeed, they have shown that nucleation can be understood by invoking proper friction relations (Dieterich, 1992; Lapusta and Rice, 2003; Uenishi and Rice, 2003). Unfortunately, it is impractical to explore the detailed physics of the nucleation event experimentally and instead, we have decided to mimic the way that triggering of rupture is achieved in numerical simulations. The triggering of natural earthquakes can be achieved either by an increase of the shear loading or by a decrease of the fault strength at a specific location. In the first case, triggering due to a loading increase can occur dynamically if the requisite jump of shear stress is provided by stress waves generated by a nearby earthquake source. It can also occur quasi-statically due to the presence of inhomogeneities or stress concentrations along a fault (e.g., kinks or forks) which, as slip accumulates, may result in a critical increment of the local resolved shear stress on the interface. In the second case, triggering occurs through a local decrease of fault strength which may result, among other things, from the flow of pore fluid into the fault interstice generating a local release of fault pressure. Both mechanisms (prestress and pressure release) have been applied in numerical simulations of earthquake rupture dynamics (Aagaard et al., 2001; Andrews, 1976; Andrews and Ben-Zion, 1997; Cochard and Rice, 2000; Fukuyama and Madariaga, 1998).

In the laboratory earthquake models described in this chapter, the triggering mechanism is a local pressure release, which is achieved by the exploding wire technique shown in **Figure 2**. A capacitor is charged by a high-voltage power supply. The charging time is determined by the resistance of the charging resistor and the capacity of the capacitor. Upon closing the switch, the electric energy stored in the capacitor causes a high current in a thin metal wire (buried inside a hole of 0.1 mm in diameter in the interface) for a short duration. The high current turns the metal wire into high-pressure,

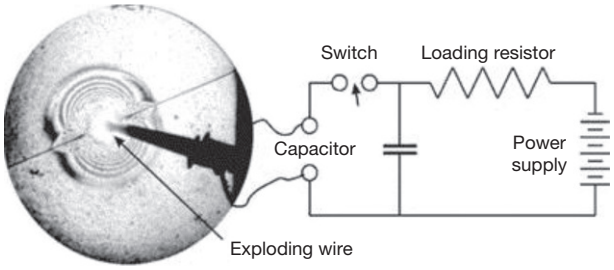


Figure 2 Schematic drawing of the exploding wire system coupled with a photoelastic fault model. Isochromatic fringes due to the explosion are visible.

high-temperature plasma in approximately 10 μs. The expansion of the high-temperature, high-pressure plasma causes a local pressure release. The adjustable power supply can provide electric potential in a wide range and different intensities and time durations of explosion can be obtained easily.

Before the explosion, the shear traction along the fault is less than the maximum static frictional strength. At the simulated hypocenter, after the explosion, the local normal traction on the fault is reduced and so is the static frictional strength. As a result, the applied shear traction, which is initially smaller than the static frictional strength and unaffected by the isotropic explosion, can be momentarily larger than the reduced frictional strength. The resulting net driving force, defined by the difference between the shear traction and the frictional resistance, drives the slip along the interface. Furthermore, the slip or the slip velocity also reduces the coefficient of friction as described by either a slip-weakening, or a velocity-weakening friction law. In other words, the frictional strength changes from static to dynamic. If the original shear traction is larger than the dynamic friction, the slip continues to propagate away from the explosion site (corresponding to the hypocenter of an earthquake) where normal traction reduction due to the explosion is not important any more. In this way a spontaneous rupture or a laboratory earthquake is triggered.

4.08.2.2.3 Choosing appropriate model materials and diagnostics

The diagnostic method of choice is dynamic photoelasticity. This technique is a classical method which measures the maximum shear stress in transparent, birefringent plates (Dally and Riley, 1991). As such the technique is very well suited for the study of shear-dominated processes. Two photoelastic solids, namely Homalite-100 and Polycarbonate, are used in all of the experiments described in the various sections of this chapter. The two materials have been chosen because of their enhanced birefringence relative to other transparent polymers such as polymethyl methacrylate (PMMA). Relevant properties of several photoelastic materials are listed in Table 1.

The stress fringe values quoted in the table are for green light at a wavelength 525 nm. The static elastic properties listed in Table 1 are from the literature (Dally and Riley, 1991), while the dynamic elastic properties (wave speeds) were measured using 5 MHz ultrasonic transducers.

A typical dynamic photoelasticity setup is shown in Figure 3. A polarized laser provides a high-intensity beam continuously at a power level of a few watts. The beam is

Table 1 Summary of optical and mechanical properties of Homalite-100 and Polycarbonate

Material property	*Homalite-100*	*Polycarbonate*
Young's modulus, E (MPa)	3860	2480
Poisson's ratio, ν	0.35	0.38
Stress fringe value, f_σ (kN m^{-1})	23.6	7.0
Yielding stress, σ_Y (MPa)	48.3	34.5
P-wave speed, C_P (km s^{-1})	2.498	2.182
S-wave speed, C_S (km s^{-1})	1.200	0.960
Density, ρ (kg m^{-3})	1230	1129

then expanded by a collimator to a size of 100 or 130 mm in diameter. The large beam is transmitted through a combination of circular polarizers and the transparent photoelastic specimen, and it is arranged so that an isochromatic fringe pattern is obtained and focused into the camera. The isochromatic fringe pattern obtained from the two-dimensional photoelastic model gives fringes along which the in-plane principal stress difference $\sigma_1 - \sigma_2$ is equal to a constant. When the fringe order N is known, the in-plane principal stress difference can be computed as follows:

$$\sigma_1 - \sigma_2 = N f_\sigma / h \qquad [3]$$

where h is the plate thickness of the model material. The high-speed camera (Cordin 220) used can take pictures at speeds up to 10^8 frames per second. In most experiments an interframe time of 2–4 μs was used.

Another and perhaps the most important reason of using brittle polymers in our experiments is the small rupture nucleation length, L_C, resulting when such materials are frictionally held together by loads on the order of 10–20 MPa. As one can see by using eqn [1] in Section 4.08.3.3, these lengths are on the order of 5 mm and unlike rocks allow for relatively small laboratory specimen sizes. If rock samples were to be used instead of stiff polymers, the resulting rupture nucleation lengths would be approximately 30 times larger under similar far-field loading and surface finish. Finally, both Homalite and Polycarbonate are very brittle and are known to deform elastically at strain rates exceeding 10^2 s^{-1}.

An example of a dynamic photoelastic pattern corresponding to a 'coherent' (glued) and horizontal ($\alpha = 0$) interface subjected to far-field tension is shown in Figure 4. This configuration which is of relevance to dynamic self-similar fracture of the mode-I (opening) type has been discussed in the recent work by Xia et al. (2006).

4.08.3 Supershear and Sub-Rayleigh to Supershear Transition in Homogeneous Fault Systems

The question of whether natural earthquake ruptures can ever propagate at supershear speeds was until recently a subject of active debate within the seismological community. This is because of the often insufficient field data, as well as the limited resolution and nonuniqueness of the inversion process. However, a recent experimental study (Xia et al., 2004) has almost settled this question by demonstrating that under

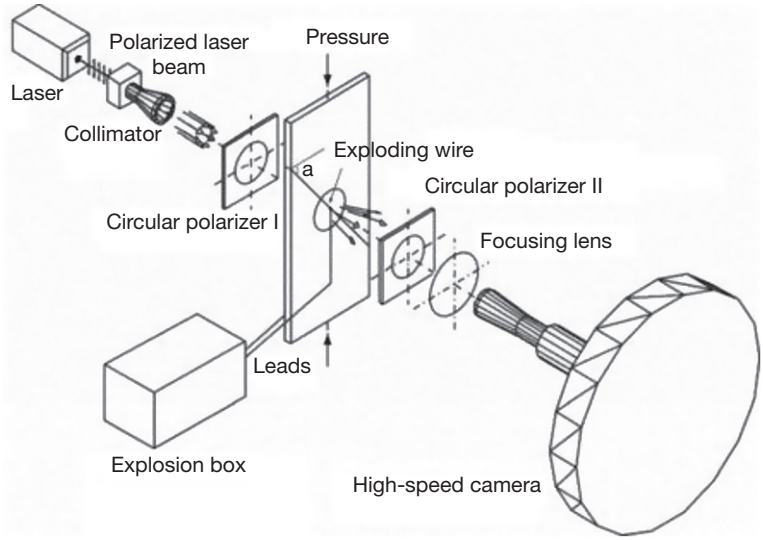


Figure 3 The setup of dynamic photoelasticity combined with high-speed photography for laboratory earthquake studies.

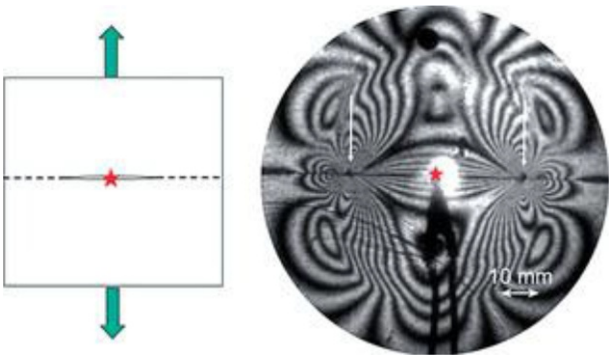


Figure 4 Specimen configuration and photoelastic fringe pattern for self-similar crack growth of a mode-I (opening) crack along a coherent interface. White arrows indicate crack tips.

controlled laboratory conditions supershear frictional rupture can occur under conditions similar to those present in some natural settings. A widespread view in seismology speaks of crustal earthquake ruptures mainly propagating at sub-Rayleigh speeds between 0.75 and 0.95 C_R (Kanamori, 1994). However, the multiplicity of independently collected evidence warrants further investigations of the mechanics of supershear rupture (velocity faster than the shear wave speed of the rock, C_S) propagation. Whether and how supershear rupture occurs during earthquakes has an important implication for seismic hazard because the rupture velocity has a profound influence on the character of near-field ground motions (Aagaard and Heaton, 2004).

In the 8 August 2003 issue of *Science*, Bouchon and Vallee (2003) took advantage of the unusual length of the rupture event and used both seismic waves and geologically observed total slip distribution to infer the rupture velocity history of the great Ms 8.1 (Mw 7.8) central Kunlunshan earthquake that occurred in Tibet on 14 November 2001. This is an extraordinary event from the point of view of both earthquake dynamics and dynamic rupture mechanics. The rupture occurred over a very long, near-vertical, strike-slip fault segment of the active Kunlunshan fault and featured an exceptionally long (400 km) surface rupture zone and large surface slip (Lin et al., 2002). Although it may not be unique, their modeling suggests speeds that are close to the Rayleigh wave speed of the crust, C_R, for the first 100 km of rupture growth, and transitioning to a supershear speed for the remaining 300 km of propagation. Their results were later corroborated by Das (private communication).

Recently, several other seismological reports also point to the possibility of supershear ruptures. These reports concentrate on events such as the 1979 Imperial Valley earthquake (Archuleta, 1984; Spudich and Cranswick, 1984), the 1992 Landers earthquake (Olsen et al., 1997), and most recently the 2002 Denali earthquake in Alaska (Ellsworth et al., 2004). The 1999 Izmit earthquake in Turkey (Bouchon et al., 2001) is yet another event featuring a very long segment of supershear rupture growth. It should be noted here that for most of those examples mentioned above (Izmit and Denali excluded), the supershear ruptures occurred only on short patches along the whole rupture length and the results are not conclusive. Bouchon and Vallee's (2003) paper together with Das's work on Kunlunshan, as well as the work of Ellsworth et al. (2004) on Denali, is the most recent of a series of investigations reporting supershear rupture growth occurring during 'large' earthquake events. Moreover, they present the first seismological evidence for transition from sub-Rayleigh to supershear.

Classical dynamic fracture theories of growing shear cracks have many similarities to the earthquake rupture processes (Broberg, 1999; Freund, 1990; Rosakis, 2002). Such theories treat the rupture front as a distinct point (sharp tip crack) of stress singularity. These conditions are close to reality in cases that feature strong 'coherent' interfaces of finite intrinsic strength and toughness. The singular approach ultimately predicts that dynamic shear fracture is allowed to propagate either at a sub-Rayleigh wave speed or at only one supershear speed,

which is $\sqrt{2}$ times the shear wave speed. As a result, it excludes the possibility of a smooth transition of a steady-state rupture from sub-Rayleigh to supershear speed for a steady-state rupture. The introduction of a distributed rupture process zone has allowed fracture mechanics to better approximate the conditions that exist during real earthquake events (Ida, 1972; Palmer and Rice, 1973). Based on this so-called cohesive zone fracture mode, there is a forbidden speed range between C_R and C_S (Burridge et al., 1979; Samudrala et al., 2002a,b). In the sub-Rayleigh speed range all speeds are admissible, but only the Rayleigh wave speed is a stable speed; in the supershear speed range all speeds are admissible, but only speeds larger than $\sqrt{2}C_S$ are stable (Samudrala et al., 2002b). Ruptures with unstable speeds will accelerate to a stable speed as determined by the fault strength and the loading conditions. The theoretical results of the cohesive zone rupture model ultimately predict that earthquake ruptures can propagate either at Rayleigh wave speed or at supershear speeds lying within the interval between $\sqrt{2}C_S$ and C_P.

Early theoretical results by Burridge and coworkers (Burridge, 1973; Burridge et al., 1979), along with numerical results by Andrews (1976) and Das and Aki (1977), have predicted the possibility of supershear rupture and have alluded to a mechanism (Rosakis, 2002) for transition from the sub-Rayleigh to the supershear rupture velocity regime. This is often referred to as the 'Burridge–Andrews' mechanism. According to the 2-D Burridge–Andrews mechanism, a shear rupture accelerates to a speed very close to C_R soon after its initiation. A peak in shear stress is found sitting at the shear wave front and is observed to increase its magnitude as the main rupture velocity approaches C_R. At that point, the shear stress peak may become strong enough to promote the nucleation of a secondary microrupture whose leading edge propagates at a supershear speed. Shortly thereafter, the two ruptures join up and the combination propagates at a speed close to C_P, the longitudinal wave speed of the solids. It is interesting that this transition was also clearly visualized by recent 2-D, atomistic calculations (Abraham and Gao, 2000; Gao et al., 2001) of shear rupture in the microscale, which provided an impressive demonstration of the length-scale persistence (14 orders of magnitude difference between the nanoscale and the scale of natural earthquake rupture) of this sub-Rayleigh to supershear rupture transition mechanism. The Burridge–Andrews mechanism is also known as the mother–daughter mechanism in mechanics literature.

For mixed-mode (tensile and shear) ruptures, a different transition model has also been suggested (Geubelle and Kubair, 2001; Kubair et al., 2002, 2003). Based on numerical simulation, they suggest that a mix-mode rupture can speed up and cross the forbidden speed range between C_R and C_S continuously. Finally, recent numerical investigations of frictional rupture have identified alternate, asperity based, mechanisms that provide a 3-D rationalization of such a transition (Day, 1982; Dunham et al., 2003; Madariaga and Olsen, 2000). In this case, 3-D effects play an important role in the transition. The rupture front focusing effect provides extra driving force to speed up the spontaneous rupture in the supershear range.

The experimental confirmation of the possibility of supershear (intersonic) fracture followed many years after the first theoretical predictions. Indeed, the long series of experiments summarized in a recent review by Rosakis (2002) demonstrate that intersonic crack growth in constitutively homogenous systems featuring coherent interfaces (interfaces with inherent strength) is possible and may also occur in various combinations of bimaterial systems. However, in all of the various cases discussed by Rosakis (2002), the cracks were nucleated directly into the intersonic regime and there was no observation of a transition from sub-Rayleigh to supershear speeds. This was due to the nature of the impact-induced stress wave loading which was imparted in the absence of preexisting static loading and the nature of the relatively strong coherence of the interface (provided by adhesive). The major differences between the conditions during earthquake rupture and those early fracture experiments have left questions regarding the plausibility of spontaneously generated intersonic rupture in frictionally held, incoherent interfaces unanswered. The recent work by Xia et al. (2004) on frictionally held interfaces has addressed this issue in detail. In the following sections, we will summarize their most important findings and we will attempt to relate these findings to field measurements related to recent natural earthquake events.

4.08.3.1 Purely Sub-Rayleigh and Purely Supershear Earthquake Ruptures

In this section, experimental parameter ranges leading to either exclusively sub-Rayleigh or exclusively supershear ruptures will be discussed. The configuration of interest is the one described in **Section 4.08.2.2**. The physics governing possible transitions between these two speed regimes within a single rupture event will be examined later. **Figure 5** shows two different experiments featuring sub-Rayleigh speed ruptures. In all the photographs, we can clearly see the circular shear wave front emitted from the simulated hypocenter. Rupture tips, characterized by the stress concentration (fringe concentrations in the photographs), are identified just behind the shear wave front. The P-wave is long gone out of the field of view (100 mm). The ruptures as shown in the photograph feature right lateral slip while the rupture tips are equidistant from the hypocenter (symmetric bilateral slip). From the rupture length history, it is possible to estimate the rupture velocity. In these two cases, the rupture velocities are very close to the Rayleigh wave speed of the material.

All experiments featuring lower inclination angles α and lower magnitude of uniaxial compression pressure P involve exactly the same features and ruptures speeds. This observation is consistent with many observations of earthquake rupture velocities in the field. As we have discussed earlier, in the sub-Rayleigh speed range the only stable rupture velocity is the Rayleigh wave speed according to various cohesive zone models of rupture. The experiments of Xia et al. (2004) described in this section confirm this observation.

In **Figure 6**, the inclination angle was kept at 25° while the pressure was increased to 15 MPa. For comparison purposes, the same time instants (28 and 38 μs after nucleation) are displayed.

In this case, the circular traces of the shear wave are also visible and are at the same corresponding locations as in **Figure 6**. However, in front of this circle supershear disturbances (propagating to the left, marked in the photograph as

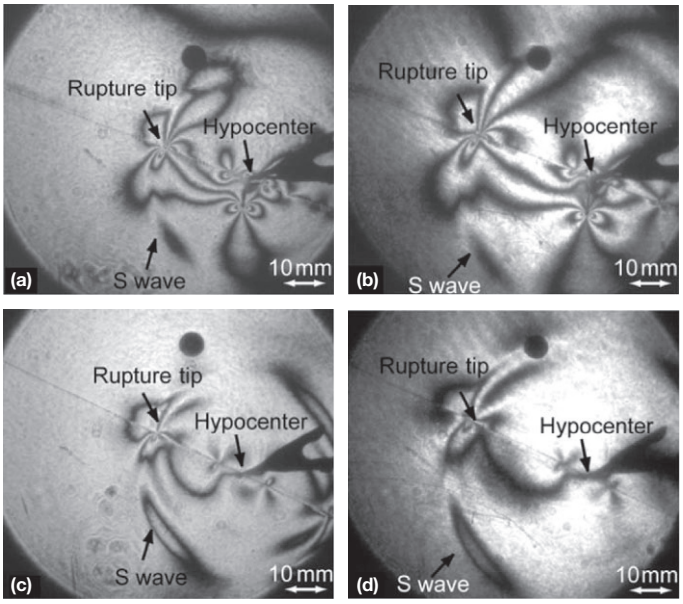


Figure 5 Experimental results of purely sub-Rayleigh cases. (a) and (b) are from one experiment with a pressure of $P=13$ MPa and angle $\alpha=20^\circ$ at time instants of 28 and 38 μs, respectively. (c) and (d) are from an experiment with a pressure of $P=7$ MPa and angle $\alpha=25^\circ$ at the time instants of 28 and 38 μs, respectively. For (a) and (b), we can also identify two mode-I cracked in the lower half of the sample caused by the explosion itself. We expect that the effect of these cracks is localized.

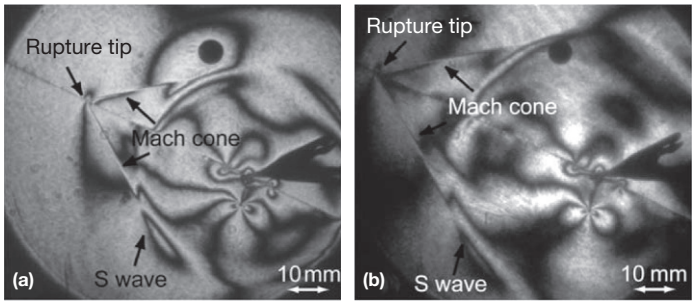


Figure 6 Experimental results for the purely supershear case. (a) and (b) both correspond to an experiment with the pressure of $P=15$ MPa and angle $\alpha=25^\circ$ at the two time instants of 28 and 38 μs, respectively.

the rupture tip and featuring a clearly visible Mach cone) are shown. The formation of the Mach cone is due to the fact that the rupture front (tip) is propagating faster than the shear wave speed of the material. For this case, the sequence of images, other than those at 28 and 38 μs, have a very similar form and reveal a disturbance that was nucleated as supershear. The speed history $v(t)$ is determined independently by either differentiating the rupture length history record or measuring the inclination angle, δ, of the shear shocks with respect to the fault plane, and using the relation $v=C_S/\sin\delta$. The speed was found to be almost constant and very close to the plane stress P-wave speed C_P of the material. These experiments provide the first evidence of supershear growth of a spontaneous shear rupture propagating along a frictional interface. The supershear rupture initiated right after the triggering of the earthquake rupture. This is determined from the fact that the Mach cones are nearly tangential to the shear wave front which was itself created at the time of rupture nucleation.

In previous experiments involving strong, coherent interfaces and impact induced stress wave loading, stable rupture velocities slightly exceeding $\sqrt{2}C_S$ were observed (Rosakis et al., 1999). This apparent discrepancy can be explained by referring to the rupture velocity dependence on the available energy per unit shear rupture advance within the supershear regime (Samudrala et al., 2002b). This dependence was obtained by means of an analytical model of a dynamic steady state, shear rupture featuring a shear cohesive zone of the velocity weakening type. The initial, static, strength of the interface was represented by the parameter τ_0. For a given normalized far-field loading level, σ_{12}^D/τ_0, this energy, shown here in normalized form (Figure 7) as a function of rupture speed, attains a maximum value at speeds closer to $\sqrt{2}C_S$ for interfaces that are 'strong.' This is consistent with the purely singular prediction which corresponds to idealized interfaces of infinite strength. To this idealized case, $(\tau_0\to\infty)$, the only value of rupture velocity corresponding to finite G is exactly $\sqrt{2}C_S$. For weaker

interfaces (or higher driving stresses), this maximum moves toward C_P. In the situation described in this section, the interface is 'weak' and the driving force (resolved shear minus dynamic frictional force) is relatively large and constant. Hence, a rupture velocity close to C_P is expected.

4.08.3.2 The Experimental Visualization of the Sub-Rayleigh to Supershear Earthquake Rupture Transition

As discussed earlier, investigating how supershear ruptures are nucleated experimentally is of great interest. For the experimental cases described in the above section, the supershear rupture event was nucleated immediately after triggering of the rupture event. Since the rupture velocity is controlled by both the inclination angle α and the magnitude of uniaxial compression P, it is possible to vary both of them carefully to suppress or perhaps delay the appearance of supershear rupture. Specifically, the inclination angle α was fixed at 25° while P was decreased in order to induce and capture the nucleation process of a supershear rupture.

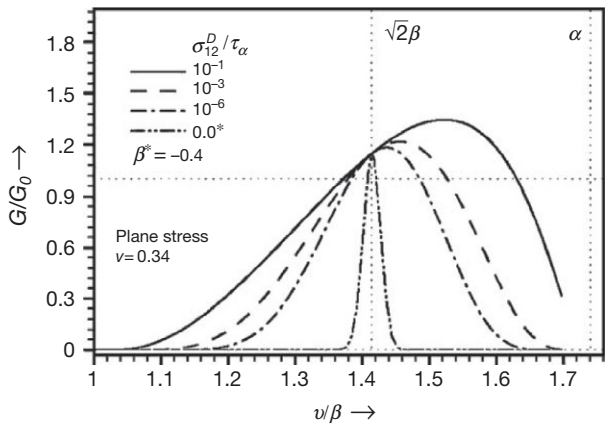


Figure 7 Normalized energy release rate as a function of intersonic rupture speed. The figure illustrates that the energy flux available to the tip has a maximum at or above $v = \sqrt{2}C_S$.

Figure 8(a)–8(c) corresponds to a case with an intermediate far-field pressure compared to the ones displayed in Figures 5(c), 5(d), and 6. Here, the angle is kept the same (25°) and the pressure is decreased to 9 MPa in an attempt to visualize a transition within our field of view (100 mm). Three different time instances of the same rupture event are displayed. In Figure 8(a), the circular traces of both P- and S-waves are visible, followed by a rupture propagating at C_R. In Figure 8(b), a small secondary rupture appears in front of the main rupture and propagates slightly ahead of the S-wave front. In Figure 8(c), the two ruptures coalesce and the leading edge of the resulting rupture grows at a speed of 1970 m s^{-1} which is very close to C_P. A well-defined 'shear' Mach cone is visible in Figure 8(c). Unlike Figure 6, the Mach lines are not tangent to the shear wave circle since the supershear rupture was not generated at initiation. Figure 8(d) displays the length versus time of the two ruptures, in which the length is directly read from the figure. The figure also compared the slopes to the characteristic wave speeds of the material before and after their coalescence. A magnified view of the secondary rupture as it nucleates in front of the main rupture is shown in the inset. Both ruptures are indicated by arrows. The transition length L for this case is approximately 20 mm, as is visually obvious by Figure 8(b).

In cases where the transition length cannot be measured directly if it is too small, an indirect method is developed. As shown in Figure 9, at time T_0, the supershear rupture is nucleated at the intersection of the shear wave front and the fault line (point A). If the transition length $L = \overline{OA}$ is very small, spatial resolution may not be good enough to measure it accurately. Alternatively, a photograph may not have been

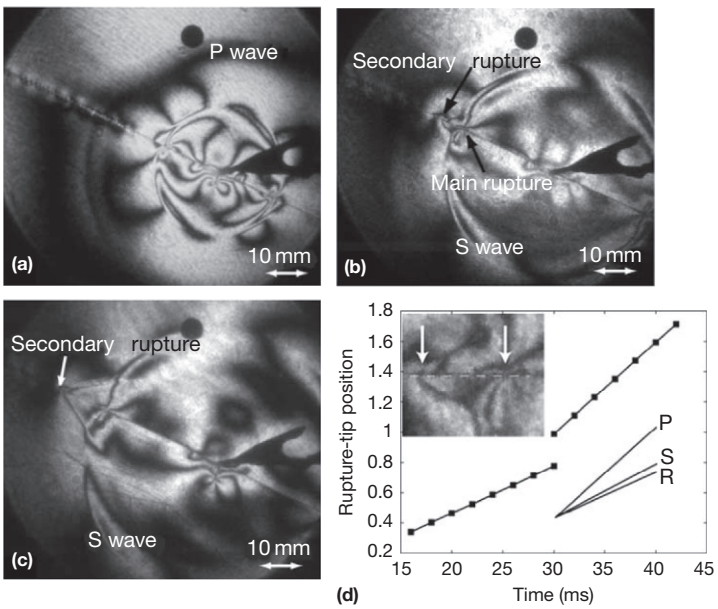


Figure 8 Visualization of the sub-Rayleigh to supershear rupture transition.

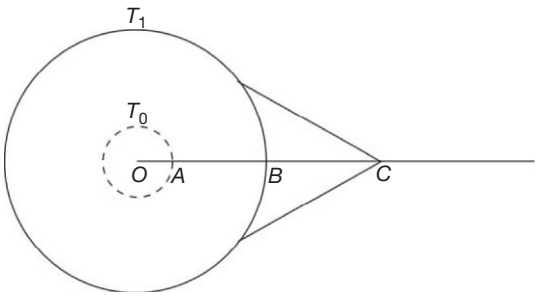


Figure 9 Method of transition length L estimation.

taken at that instant. Assuming that at all times before transition the rupture speed is very close to C_S and that the shear wave position (A) and supershear rupture-tip position (B) at a later time instance T_1 can be measured, the transition length can be inferred by pure geometry. To do so we observe that relations $\overline{OA} = C_S T_0$, $\overline{OB} = C_S T_1$, and $\overline{AC} = V(T_1 - T_0)$ hold provided that the supershear rupture tip also grows at a constant speed $V > C_S$ (as confirmed by Figure 8).

A simple manipulation gives $-AC/-OB = (1 - T_0/T_1)V/C_S$, which leads to $T_0 = [1 - (C_S/V)(-AC/-OB)]T_1$. Multiplication of both sides of the last relation by C_S gives $L = -OA = C_S T_0 = [-OB - -OC \times C_S/V]$. This relation allows for the estimation of the supershear rupture transition length, L, even if no photographic records have been obtained at or before the time instance of the transition. Several results listed in this chapter were either obtained or verified by using this method.

4.08.3.3 A Theoretical Model for the Sub-Rayleigh to Supershear Transition

The above physical picture is comparable with the Burridge–Andrews mechanisms already described in Section 4.08.1. Andrews (1976, 1985) quantified this transition in a parameter space spanned by a normalized supershear transition length L/L_C and the nondimensional driving stress parameter s [$s = (\tau^y - \tau)/(\tau - \tau^f)$]. The parameters τ, τ^y, and τ^f are the resolved shear stress on the fault, the static frictional strength, and the dynamic strength of the fault, respectively; they describe the linear slip-weakening frictional law (Ida, 1972; Palmer and Rice, 1973) used in Andrews' computations.

The Andrews result can be symbolically written as $L = L_C f(s)$. The function $f(s)$ has been given numerically by Andrews as an increasing function of s, and can be well approximated by the equation $f(s) = 9.8(1.77 - s)^{-3}$. The normalizing length L_C is the critical length for unstable rupture nucleation and is proportional to the rigidity G and to d_0, which is defined as the critical or breakdown slip of the slip-weakening model. L_C can then be expressed as

$$L_C = \frac{1+\nu}{\pi}\frac{(\tau^y - \tau^f)}{(\tau - \tau^f)}Gd_0 \qquad [4]$$

By adopting the above relation to the geometry of the experimental configuration and by using eqn [1], the transition length L is found to be inversely proportional to the applied uniaxial pressure P and to be governed by the following general functional form:

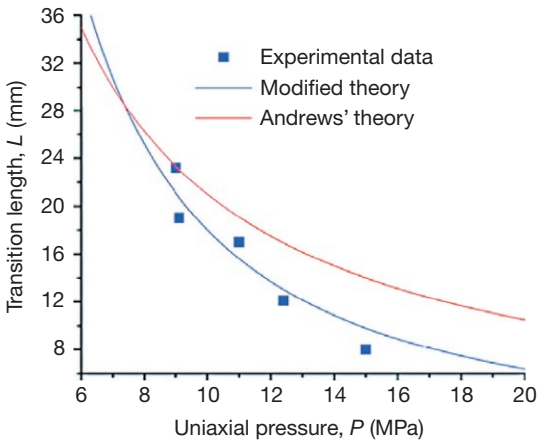


Figure 10 Transition length as a function of the far-field load.

$$L = f(s)\frac{1+\nu}{\pi}\frac{(\mu^s - \mu^d)}{(\tan\alpha - \mu^d)^2}\frac{Gd_0}{P} \qquad [5]$$

Figure 10 displays the dependence of the transition length L on pressure from a set of experiments corresponding to the same inclination angle of 25° ($s = 0.5$) and identical surface finish (roughness is about 17 μm). The static frictional coefficient was measured to be $\mu^s = 0.6$ using the traditional inclined plane method. In this method, one block was placed on top of an inclined plane and the inclination angle until the block slides. By this way the critical inclination angle Ψ_C was measured. The static coefficient of friction was then determined from the relation $\mu^s = \tan^{-1}(\Psi_C)$. To estimate the average, steady state, dynamic frictional coefficient, α was gradually increased from 10° to a critical angle α_c at which slip was initiated under the action of far-field loads and the dynamic triggering. It was assumed that the shear traction was approximately equal to the dynamic frictional strength at this critical angle α_c. This angle was found to be between 10° and 15°, from which the average coefficient of dynamic friction was estimated to be $\mu^d = 0.2$. This low value of dynamic friction coefficient is very consistent with recent experimental measurements conducted by Tullis and his research group at slip velocities approaching seismic slip rates (Di Toro et al., 2004; Tullis and Goldsby, 2003). Using $\mu^s = 0.6$ and $\mu^d = 0.2$, the experiments can be compared to the predictions of eqn [5] of Andrews' theory as shown in Figure 10. Although the theory qualitatively captures the decreasing trends of the experiments, the data exhibit a dependence on pressure that is visibly stronger than P^{-1}.

A natural way to modify Andrews' results and to introduce some scaling into the pressure dependence of L is to consider the effect of pressure on the critical breakdown slip d_0. As pointed out by Ohnaka, on the basis of frictional experiments on rocks, and as also supported by dimensional analysis, there exists a linear relation between a characteristic surface length (half-distance between contacting asperities, denoted as D in this case) and the critical slip distance d_0 as (Ohnaka, 2003)

$$d_0 = c\left[\left(\tau^y = \tau^f\right)/\tau^f\right]^M D \qquad [6]$$

where c and M are experimentally determined constants. In addition, D depends on the normal stress, σ, applied on the

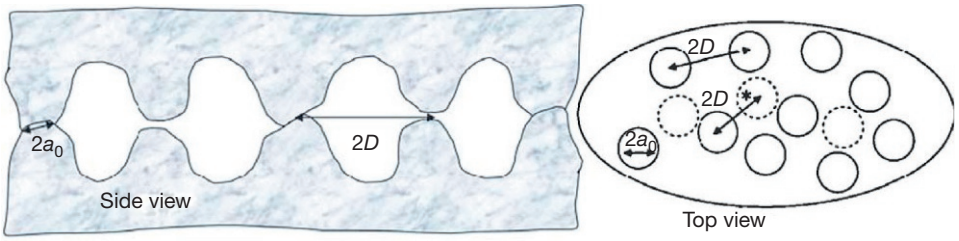


Figure 11 Schematic drawing of the microcontact-based frictional model. The top figure is the side view of the contact and the bottom figure is the top view. As the normal force increases, the number of contacts, n, increases.

fault (which is $\sigma = P\cos^2\alpha$ in this case). As shown in Figure 11, a classical plastic contact model may be used to establish this dependence. In this model the average radius of n contacting asperities is a_0 (assumed as a constant in this model). As the pressure over a macroscopic contact area A $(=n\pi D^2)$ is increased, the number of contacts, n, as well as the real contact area $A_r\left(=n\pi a_0^2\right)$, also increases.

By defining the hardness H as the ratio of the total normal force N to the real contact area A_r (Bowden and Tabor, 1986), N can be expressed as

$$N = HA_r = Hn\pi a_0^2 = \sigma A = AP\cos^2\alpha \qquad [7]$$

Substitution of A and A_r in terms of D and a_0 gives $D = \sqrt{H}a_0\cos\alpha P^{-1/2}$. Using the linear relation between D and d_0, the breakdown slip is further found to depend on the pressure as $d_0 \sim P^{-1/2}$. By substituting the above relations into the expression relating L and d_0/P, discussed above, a modified expression of transition length to pressure that features a stronger dependence $(L \sim P^{-3/2})$ on pressure emerges. For the loading conditions and geometry of the experimental configuration presented in Figure 1, this expression becomes

$$L = f(s)\frac{1+\nu}{\pi}G\frac{\mu^s-\mu^d}{\left(\sin\alpha-\mu^d\cos\alpha\right)^2}2c\left(\frac{\mu^s-\mu^d}{\mu^s}\right)^M \times\sqrt{H}a_0P^{-3/2}\cos^{-1}\alpha \qquad [8]$$

As shown in Figure 10, this modified relation agrees very well with the experimental data presented in this chapter for appropriate choices of material and geometrical parameters of the micromechanics contact model. In plotting Figure 10, the hardness and the elastic constants of Homalite were used while the friction coefficients were taken to be $\mu^s = 0.6$ and $\mu^d = 0.2$ as estimated above. The values for the constants c and μ were from the experiments of Ohnaka (2003). The asperity radius a_0 was taken to be equal to 1.78 μm, for a best fit to the experimental data. This value was found to be consistent to averaged measurements of asperity radii obtained through microscopic observations of the sliding surfaces. Relation [8] can also easily be generalized to biaxial loading conditions. Indeed, if $P_1 = P$ and $P_2 = bP$ denote the horizontal and vertical components of the applied loads such a situation, the transition length can be expressed as

$$L = f[s(\alpha)]\frac{1+\nu}{\pi}G\frac{2c\left(\mu^s-\mu^d\right)^{1+M}\left(\cos^2\alpha+b\sin^2\alpha\right)^{1.5}}{\left[(1-b)\sin\alpha\cos\alpha-\mu^d\left(\cos^2\alpha+b\sin^2\alpha\right]^2} \times\left(\mu^s\right)^{-M}\sqrt{H}a_0p^{-1.5} \qquad [9]$$

In addition to tectonic loading magnitude and direction relative to fault inclination, eqn [9] involves the material properties G, ν, and H, the frictional properties μ^s and μ^d, as well as fault-related length scale a_0. Because of this, it can perhaps be used as a simple scaling relation connecting laboratory observations to reports of supershear transition in 'large' natural earthquake rupture events such as these related to 2001 Kunlushan event in Tibet.

For making order of magnitude-level contact with seismological applications, we rewrite the general form of eqn [5] in terms of the effective stress $\tau^e = \tau - \tau^f$, a commonly used and estimated seismological parameter. Equation [5] can be rewritten as

$$L = f(s)\frac{1+\nu}{\pi}(1+s)\frac{Gd_0}{\tau^e} \qquad [10]$$

Application of this equation to both seismic faulting and laboratory data allows us to scale the transition length L from laboratory to seismological conditions. The effective stress τ^e in our experiment is chosen to be of the same order of magnitude as that measured in seismology. The ratio of rigidity of the Earth's crust to Homalite is about 25. From the experiment described in Figure 8 ($P = 9$ MPa and $\alpha = 25°$) the transition length is estimated to be $L = 20$ mm from which $d_0 = 10$ μm is obtained using eqn [4]. The values of d_0 for large earthquakes are often estimated as 50 cm to 1 m (Ide and Takeo, 1997). If s is approximately the same under laboratory and crustal conditions, the transition length for earthquakes can be estimated to be in the range between 25 and 50 km. Because s can be different and the estimate of d_0 for earthquakes is very uncertain at present, this value should only be taken as an order of magnitude estimate. Nevertheless, it is of the same order as that inferred for Kunlunshan (100 km) and for Denali (75 km). The large transition length required for supershear is perhaps one of the reasons that relatively few earthquake events have been observed to feature such high rupture velocities and that all of them correspond to large magnitude earthquakes. If the tectonic stress is well below the static fault strength (i.e., large s), the transition length becomes too large for many earthquake ruptures to attain supershear. The observation that during several large earthquakes the rupture velocity became very fast, possibly supershear, suggests that the tectonic stress is fairly close to the static fault strength (i.e., small s), which has important implications for the evolution of rupture in large earthquakes.

4.08.3.4 Discussion

A final word of caution is in order here. It has recently been shown in recent preliminary numerical simulations of the homogeneous experiments described in this section

(Lu et al., 2005) that the detailed variation of the transition length on the load P may also strongly depend on the nature of the rupture nucleation process. The severity of this effect is still unknown. If this dependence proves to be significant, an experimental investigation involving various pressure release histories (intensities, rise times, and durations) will be necessary in order to complete the experimental investigation of supershear transition presented above. A related issue regarding the influence of nucleation conditions in a bimaterial setting (see **Section 4.08.4**) was recently discussed by Shi and Ben-Zion (2006).

4.08.4 Directionality of Ruptures Along Faults Separating Weakly Dissimilar Materials: Supershear and Generalized Rayleigh Wave Speed Ruptures

In most mature fault systems, the elastic properties vary across the fault plane (Magistrale and Sanders, 1995; Peltzer et al., 1999) and also the shear wave speeds may vary by as much as 30% (Ben-Zion and Huang, 2002; Cochard and Rice, 2000). Recently, Rubin and Gillard (2000) studied several thousands of pairs of consecutive earthquakes that occurred on a segment of the central San Andreas fault, south of the Loma Prieta rupture. Among the second events of each pair, they found that over 70% more occurred to the northwest than to the southwest. They interpret this asymmetry as being a result of the contrast in material properties across the fault. Indeed, at this location of the San Andreas Fault, the rock body is more compliant northeast of the fault than it is southwest (Eberhart-Phillips and Michael, 1998).

Early theoretical and numerical studies of rupture that employ various frictional laws with a constant coefficient of friction which would be independent of slip or slip rate (Adams, 1995; Andrews and Ben-Zion, 1997; Cochard and Rice, 2000; Harris and Day, 1997; Heaton, 1990; Ranjith and Rice, 1999; Rice et al., 2001; Weertman, 1980) imply that if rupture occurs on the boundary between two frictionally held solids, having different elastic properties and wave speeds, such a rupture may preferentially propagate in a particular direction. This is the same direction as the direction of slip in the lower wave speed solid. This direction is often referred to as the 'preferred' direction (Ben-Zion, 2001). Since such implied directionality of fault rupture may have a profound influence on the distribution of damage caused by earthquake ground motion, it would be extremely useful if this behavior could be confirmed under controlled laboratory conditions. While many of the physical aspects of dynamic rupture (including supershear) are recently becoming progressively clearer in relation to homogeneous faults (Ben-Zion, 2001; Rice, 2001; Rosakis, 2002; Xia et al., 2004), the behavior of spontaneously nucleated ruptures in inhomogeneous faults, separating materials with different wave speeds, has until recently (Xia et al., 2005) remained experimentally unexplored. In this section we elaborate on the results of this experimental investigation and we discuss their implications in relation to the theoretical concept of a 'preferred' rupture direction in the presence of bimaterial contrast.

The recent large earthquakes (1999 Izmit and Düzce) and the seismic migration history along the North Anatolian Fault may represent a unique field example of the effect of the material contrast across the fault. The 1999 Izmit and Düzce events featured both supershear and sub-Rayleigh rupture branches (Bouchon et al., 2001). Most significantly, they are the last of a series of large ($M \geq 6.8$) earthquakes that have occurred since 1934 in the North Anatolian Fault. These earthquakes have occurred on a rather long and allegedly inhomogeneous fault system (Le Pichon et al., 2003; Zor et al., 2006) that has hosted tens of major migrating earthquakes in the past century. Following the work of Stein et al. (1997) and of Parsons et al. (2000), tens of large ($M \geq 6.8$) earthquakes occurred over 1000 km along the North Anatolian Fault between the 1939 earthquake at Ercinzan and the 1999 Izmit and Düzce earthquakes. Such a long series of earthquakes are believed to be textbook examples of how the transfer of stress from a recent nearby event can trigger the next major event in due time. This presumably happens by adding or transferring stress to the fault segment, which is adjacent to the tips of a segment that has last failed. The stress distribution is highly nonuniform and it occurs in addition to the long-term stress renewal and to the preexisting stress inhomogeneities. However, as much as this model seems to be complete and convincing, a few questions remain that need to be resolved and are of relevance to the work described here.

4.08.4.1 Two Types of Ruptures Along Inhomogeneous Fault Systems

Inhomogeneous faults separate materials with different wave speeds. When such faults experience spontaneous rupture, the equibilateral symmetry, expected in the homogeneous case, is broken. This leads to various forms and degrees of rupture directionality. Dynamic rupture along bimaterial interfaces is known to involve substantial coupling between slip and normal stress (Ben-Zion, 2001; Rice, 2001; Weertman, 1980). As a consequence, the relative ease or difficulty for a rupture to propagate in a specific direction along a bimaterial interface should be closely related to the degree of mismatch in wave speeds. This phenomenon is also related to the fault's frictional characteristics which play a dominant role in facilitating the phenomenon of normal shear coupling. For bimaterial contrast with approximately less than 35% difference in shear wave speeds (as in the case of most natural faults), generalized Rayleigh waves can be sustained. These waves are waves of frictionless contact propagating at a speed, C_{GR}, called the generalized Rayleigh wave speed (Rice, 2001).

The 1980 rupture solution by Weertman (1980) involves a dislocation like sliding pulse propagating subsonically with a velocity equal to C_{GR} along an interface governed by Amonton–Coulomb friction. However, the classical Amonton–Coulomb's description has been shown to be inadequate for addressing fundamental issues of sliding (Ranjith and Rice, 2001), since sliding becomes unstable to periodic perturbations. Instability, in the above sense, implies that periodic perturbations to steady sliding grow unbounded for a wide range of frictional coefficient and bimaterial properties (Adams, 1995; Renardy, 1992). The growth rate is proportional to the wave number. This issue has been discussed in relation of both homogeneous and bimaterial systems (Rice, 2001). For bimaterial systems and in particular, when generalized

Rayleigh waves exist, Ranjith and Rice (2001) demonstrate that unstable periodic modes of sliding appear for all values of the friction coefficient. Mathematically, instability to periodic perturbations renders the response of a material interface to be ill-posed (no solution exists to the problem of growth of generic, rather than periodic, self-sustained perturbations to steady sliding). The problem is regularized by utilizing an experimentally based frictional law (Prakash and Clifton, 1993), in which shear strength in response to an abrupt change in normal stress evolves continuously with time (Cochard and Rice, 2000; Ranjith and Rice, 2001). In such a case, the problem becomes well-posed and generic self-sustained pulse solutions exist while numerical convergence through grid size reduction is achieved (Cochard and Rice, 2000; Coker et al., 2005). However, despite the fact that this special frictional law provides regularization, self-sustained slip pulses may still grow in magnitude with time. This is a phenomenon that has been demonstrated numerically by Ben-Zion and Huang (2002). Moreover, self-sustained pulses were found to exist and to propagate at discrete steady velocities and at specific directions along the inhomogeneous interface by analytical (Ranjith and Rice, 2001) and numerical means (Andrews and Ben-Zion, 1997; Cochard and Rice, 2000).

Two types of such steady, self-sustained pulses were discovered by Ranjith and Rice (2001) theoretically. Consistent with Weertman's analysis (Weertman, 1980), the first type corresponds to rupture growth in the direction of sliding of the lower wave speed material of the system. This direction is referred to in the literature (Ben-Zion, 2001; Rice, 2001) as the 'positive' direction and sometimes as the 'preferred' direction (Ben-Zion, 2001). The rupture pulses belonging to this type are subshear and always propagate with a steady velocity $V = +C_{GR}$, where the plus denotes growth in the 'positive' direction. Thus, in this work these rupture pulses will be referred to as 'positive' generalized Rayleigh ruptures and will be abbreviated as '+GR' ruptures. The second type of self-sustained rupture corresponds to growth in the direction opposite to that of sliding in the lower wave speed material of the bimaterial system. This direction is often referred to as the 'negative' direction or 'opposite' direction (Cochard and Rice, 2000). Such ruptures are supershear and they always propagate with a steady velocity that is slightly lower than the P-wave speed of the material with the lesser wave speed $\left(V = -C_P^2\right)$. Such ruptures are generated for sufficiently high values of the coefficient of friction (Ranjith and Rice, 2001) and are less unstable than the '+GR' ruptures described above (Cochard and Rice, 2000). In this chapter such ruptures will be abbreviated as '$-P_{SLOW}$' ruptures. This second type of rupture was later studied by Adams (2001), who showed that the leading edges of these supershear (intersonic) ruptures are weakly singular, a result which is consistent with numerical analysis (Cochard and Rice, 2000).

From the point of view of numerics, the early work of Andrews and Ben-Zion (1997) has brought to light the persistence and interesting properties of rupture pulses of the '+GR' type. This was possible even in the ill-posed context of sliding governed by Amontons–Coulomb friction before much of the theoretical concepts were at hand. In their work, the sliding '+GR' pulses were triggered by a local release of interfacial pressure spread out over a finite region at the interface and over finite time. No pulses of the second type ('$-P_{SLOW}$' pulses) were excited in these early simulations despite the fact that the coefficient of friction was high enough to have permitted their existence as suggested by the modal analysis of Ranjith and Rice (Cochard and Rice, 2000). Perhaps this is not altogether surprising if one considers the fact that in these early simulations, nucleation was achieved by a 'biased' localized stress drop having a favored propagation direction. In particular, the 'positive' direction of growth was artificially seeded in the nucleation process (Ben-Zion, 2006). The subsequent numerical simulations of Cochard and Rice (2000), which utilized the modified Prakash and Clifton law, still featuring a constant coefficient of friction, were able to sequentially excite regularized self-sustained pulses of both types. This was achieved by introducing small changes in the parameters of the friction law and in the geometry of the nucleation zone. At the same time, no simultaneous excitation of both modes was reported. Moreover, the '$-P_{SLOW}$' pulses were found to be slightly more difficult to excite than the '+GR' pulses. However, the degree of relative difficulty was not examined in detail. In partial agreement to the above 2-D numerical studies, Harris and Day (1997) demonstrated the simultaneous existence of both types of sliding modes, propagating in opposite directions during the same rupture event. They considered various bimaterial and trilayered configurations featuring modest wave speed mismatch and a slip-weakening frictional law. As first pointed out by Xia et al. (2005), the inconsistency between the various numerical studies is most probably due to the different friction laws utilized. Indeed, up to very recently, all studies except this of Harris and Day (1997) have assumed a constant coefficient of friction. The need for experimental analysis becomes clear at this point since only experiments can be used to judge the physical relevance of various assumed friction laws and to validate various proposed numerical methodologies. As emphasized by Xia et al. (2005), the goal of some of the early theoretical and numerical studies (Adams, 1995, 1998; Andrews and Ben-Zion, 1997; Cochard and Rice, 2000; Ranjith and Rice, 2001; Weertman, 1980) was to investigate what kind of unstable slip would develop on a surface which, as judged from conventional friction notions, was superficially stable, in the sense that its friction coefficient, μ, did not decrease, or vary otherwise, with slip and/or slip rate. For most brittle solids, however, ample evidence exists that, μ does decrease with increase of slip and/or slip rate (or, more fundamentally, μ varies with slip rate and contact state). As a result, a proper model for natural faulting along a bimaterial interface should include both a weakening of μ and the slip-normal stress coupling effects of the bimaterial situation. Indeed, various models for μ are expected to strongly influence the effectiveness of the bimaterial contrast in either enhancing or retarding rupture growth. Such a weakening model was first included by Harris and Day (1997) in a bimaterial context. Given the above, Xia et al. (2005) first pointed out that it would be an invalid interpretation of the results of the earlier set of papers (Adams, 1995, 1998; Andrews and Ben-Zion, 1997; Cochard and Rice, 2000; Ranjith and Rice, 2001; Weertman, 1980) to conclude that the rupture scenarios (including preference for specific rupture mode) which they predict constitute the full set of possibilities available to a real earthquake, of which μ decreases with increasing slip and/or

slip rate. Indeed, according to Xia et al. (2005) the consistently bilateral nature of rupture predicted by Harris and Day (1997) is an indication of the effect of including a strongly slip-weakening frictional law in their calculations. In addition, recent refined calculations by Harris and Day (2005) have reconfirmed their conclusions within a fully 3-D setting. These conclusions, however, have recently been partially challenged by Shi and Ben-Zion (2006) on the basis of an extensive parameter study. A more complete analysis of this ongoing discussion will be presented in **Section 4.08.4.5**.

In the immediately following section, we elaborate on recent experiments (Xia et al., 2005) designed to investigate some of the issues discussed above. We also briefly summarize the recent debate on the existence or nonexistence of a preferable rupture direction which has been triggered by these experiments.

4.08.4.2 Experimental Setup

The experiments described in this section mimic natural earthquake rupture processes in fault systems, where bimaterial contrast between intact rock masses seldom feature more than a 30% difference in shear wave speeds (Rice, 2001).

The laboratory fault model is similar to the one described in **Figure 1(a)** and is shown in **Figure 12**. The figure shows a Homalite-100 plate (material 1, top) and a Polycarbonate plate (material 2, bottom) that are held together by far-field load, P. The higher wave speed material at the top (Homalite-100) has a shear wave speed $C_S^1 = 1200\,\mathrm{m\,s^{-1}}$ and a longitudinal wave speed $C_P^1 = 2498\,\mathrm{m\,s^{-1}}$. The lower wave speed material at the bottom (Polycarbonate) has a shear wave speed $C_S^2 = 960\,\mathrm{m\,s^{-1}}$ and a longitudinal wave speed $C_P^2 = 2182\,\mathrm{m\,s^{-1}}$. The fault is simulated by a frictionally held contact interface forming an angle to the applied load that is varied to mimic a wide range of tectonic load conditions. Spontaneous rupture is triggered at the hypocenter through the exploding wire mechanism described in **Chapter 4.02**. The static compressive load P is applied through a hydraulic press. By arbitrary convention, the fault line runs in the east–west direction with the lower wave speed solid located at the south side. This choice was motivated in order to eventually facilitate comparisons with the 1999 Izmit earthquake in Turkey. According to the work of Le Pichon et al. (2003), the material to the south of the North Anatolian Fault (at its western end near the sea of Marmara) is the lower wave speed solid. As viewed from the camera, a rupture will produce right lateral slip.

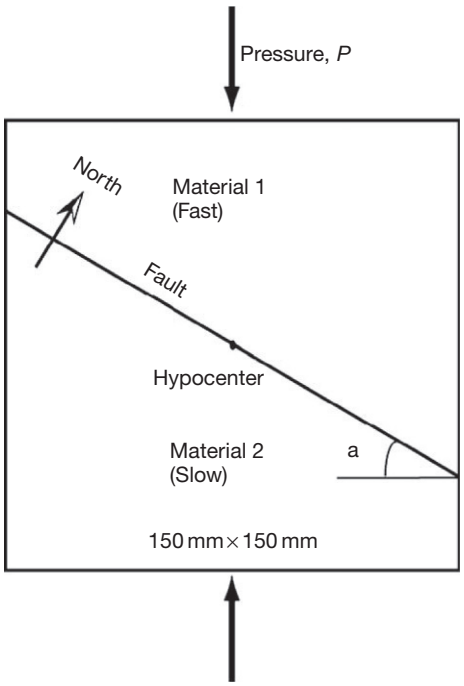


Figure 12 Laboratory earthquake fault model composed of two photoelastic plates of the same geometry.

The ratio of shear wave speeds, $C_S^1/C_S^2 = 1.25$, was chosen to be within the naturally occurring bimaterial range so that the interfacial phenomena can be applied to some field observations. In particular, the bimaterial difference is big enough to allow for a high-enough growth rate of sliding instabilities (Rice, 2001) and to permit the clear distinction between various wave speeds. Within roughly the same range generalized Rayleigh waves exist as well. The shear wave speeds were directly measured for each material by following the shear wave fronts through high-speed photography and photoelasticity. The listed P wave speeds were calculated by using measured values of Poisson's ratios ($\nu^1 = 0.35$, $\nu^2 = 0.38$) and by using the listed shear wave speeds. An independent measurement of the P-wave speeds in the plates using ultrasonic transducers confirmed these listed values to within 5%. The value of C_{GR} can be determined (Rice, 2001) from the equation:

$$f(V) = \left(1 - b_1^2\right)\alpha_1 G_2 D_2 + \left(1 - b_2^2\right)a_2 G_1 D_1 = 0$$

where

$a_n = \sqrt{1 - V^2/\left(C_P^n\right)^2}$, $b_n = \sqrt{1 - V^2/\left(C_S^n\right)^2}$, $D_n = 4a_n b_n - \left(1 - b_n^2\right)^2$, V is the rupture speed, G_n are rigidities of the two materials, and $n = 1, 2$. Substituting the material constants for Homalite-100 and Polycarbonate into the equation, the generalized Rayleigh wave speed is calculated as $C_{GR} = 950\ \mathrm{km\ s^{-1}}$. This is a value that is extremely close to the shear wave speed of Polycarbonate.

4.08.4.3 Experimental Results

More than 30 experiments featuring different angles, α (20°, 22.5°, and 25°), and far-field loading, P (10–18 MPa), were performed and the repeatability of rupture events was confirmed. The higher level of angles was limited by the static frictional characteristics of the interface. Depending on P and α, three distinct and repeatable rupture behaviors were observed. In all cases, the two separate, semicircular traces of the shear waves in the two materials were clearly visible as discontinuities in the maximum shear stress field. The ruptures were always bilateral and became progressively asymmetric with time, within the time window of all experiments. As shown in **Figure 13**, two distinct rupture tips, one moving to the west and the other moving to the east, with velocities V_E and V_W, respectively, were identified by a distinct concentration of fringe lines. For this case (case-1), both tips propagate at subshear velocities $V_E < V_W < C_S^2 < C_S^1$. Differentiation of the rupture length–time histories, obtained from a series of high-speed images, allows for the estimation of the rupture velocity histories. On the one hand, the rupture moving to the west is

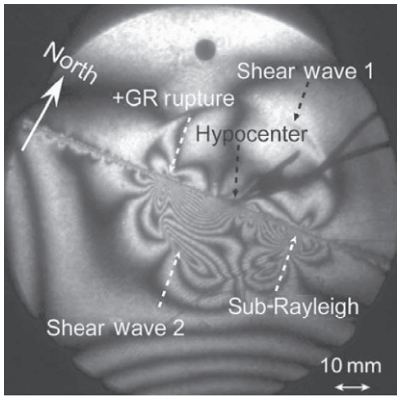


Figure 13 The photoelastic patterns for an experiment with $\alpha = 22.5$, $P = 17$ MPa, and smooth surface. Both ruptures to the east and the west are subshear (case-1).

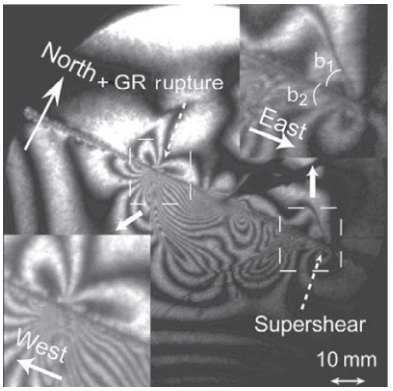


Figure 14 For $\alpha = 25°$, $P = 17$ MPa, and smooth surface finish the bilateral rupture features two distinct tips. The one moving to the west (positive direction) has a velocity $V_W \approx +C_{GR}$, while the one moving to the east (opposite direction) is supershear (case-2).

the one propagating in the direction of sliding of the lower wave speed material (positive direction). Within experimental error this rupture was found to grow at a constant velocity equal to the speed of the generalized Rayleigh waves ($V_W = 950$ m s$^{-1} \approx +C_{GR}$). The rupture moving to the east, on the other hand, was the one propagating in the direction opposite to that of sliding in the lower wave speed material (opposite direction). This rupture grew at an almost constant sub-Rayleigh velocity of $V_E = -900$ m s^{-1}, which is clearly slower than the Rayleigh wave speed, C_R^2, of the slower wave speed material. The observations were very similar for smaller angles, α, and compressive loads, P, as well. The rupture to the west (positive direction) always propagated with $+C_{GR}$. The rupture to the east remained sub-Rayleigh $\left(V_E < C_R^2 < C_R^1\right)$. However, its velocity varied across experiments with different load levels and angles. In particular, smaller angles of α (or smaller values of the s factor described earlier) and lower P resulted in V_E being lower fractions of C_R^2. Judging from the number of near-tip fringes per unit area, the eastward moving rupture resulted in a visibly smaller level of stress drop than the one moving to the west. This observation is consistent with predictions by Cochard and Rice (2000) and Adams (2001) who also predict a weaker singularity for ruptures moving in the negative direction.

A very distinct but equally repeatable rupture case (case-2) was observed for higher values of α and P. These conditions correspond to much higher values of driving stress or to conditions closer to incipient uniform sliding of the entire interface (smaller values of s). A typical example corresponding to $\alpha = 25°$ and $P = 17$ MPa is shown in Figure 14. In this case the rupture is still bilateral with a westward tip trailing behind both shear wave traces. This tip moves at a constant velocity $V_W \approx +C_{GR}$ along the 'positive' direction. This observation is identical to the situation described above in relation to lower values α and P. The eastward moving tip however is clearly different from the previously described case. Its tip is moving with a velocity faster than both the shear wave speeds. Moreover, its structure, shown in detail in the upper insert of Figure 14, is distinctly different to the structure of the sub-Rayleigh, westward moving rupture shown in the lower inset.

As conclusive proof of its supershear velocity, two distinct shear shock waves are clearly visible. The magnitude of the velocity of the eastward rupture V_E is 1920 m s^{-1}, which is approximately 12% less than the P-wave speed, C_P^2, of the lower wave speed material. V_E is also equal to 1.6 C_S^2, or is slightly higher than $\sqrt{2}C_S^1$. The upper inset in Figure 14 shows two clear lines of discontinuity in the maximum shear contours of photoelasticity. Each of these lines (shear shock waves) is located at two different angles $\beta_1 = 41°$ and $\beta_2 = 30°$, to the north and to the south of the fault respectively. The two angles β_n ($n = 1, 2$) are related to the shear wave speeds C_S^n and to the rupture velocity V_E, by $\beta_n = \sin^{-1}\left(V_E/C_S^n\right)$. This relation provides independent means of estimating V_E from each individual frame of the high-speed camera record without reliance on the less accurate rupture length history. Both methods yield consistent values of $V_E = -1920$ m s^{-1}.

Both cases described above feature westward moving ruptures that are of the '+GR' type. Irrespective of the values of α and P, these ruptures have a constant speed $V_W \approx +C_{GR}$, and they propagate in the 'positive' direction. However, those two cases also feature eastward ruptures that are distinctly different in nature. For sufficiently low P and α (or large s), the eastward ruptures, which propagate in the opposite direction, are purely sub-Rayleigh within the time window of our experiments. For large-enough P and α however, eastward ruptures propagate in the opposite direction with a constant supershear velocity, which is slightly less than C_P^2 and are thus of the '$-P_{SLOW}$' type.

To visualize an intermediate situation and a controlled transition from one case to the other within the field of view, P was reduced to 13 MPa. For this case (case-3), Figure 15 shows a smooth transition from case-1 to case-2 within the same experiment. While the westward rupture remains of the '+GR' type throughout the experiment, the eastward rupture jumps from a constant sub-Rayleigh velocity (-910 m s^{-1}) to a constant supershear velocity (-1920 m s^{-1}), and thus transitions to the '$-P_{SLOW}$' type. The rupture length plot of Figure 16 also shows the abrupt transition of the eastward rupture from a sub-Rayleigh velocity to a velocity slightly less than C_P^2. This happens at a transition length, L, which is approximately equal to 25 mm. However, the westward rupture retains its constant $+C_{GR}$ velocity throughout the experiment. The eastward transition behavior of case-3 is qualitatively similar to the one we have discussed in relation to homogeneous interfaces

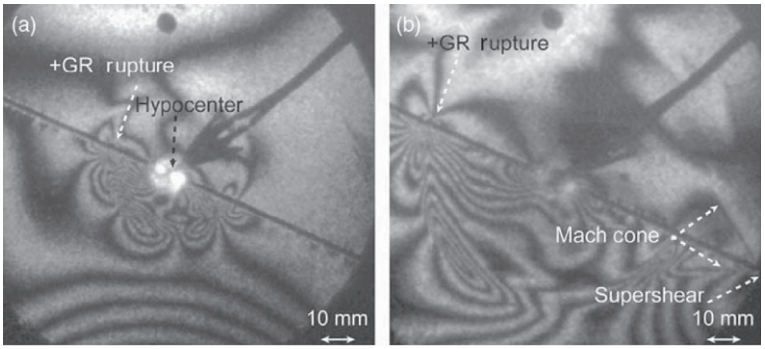


Figure 15 Experimental results for $\alpha = 25°$, $P = 13$ MPa, and rough surface showing transition of the eastward moving rupture to supershear. The westward rupture retains a constant velocity $V_W \approx +C_{GR}$ (case-3).

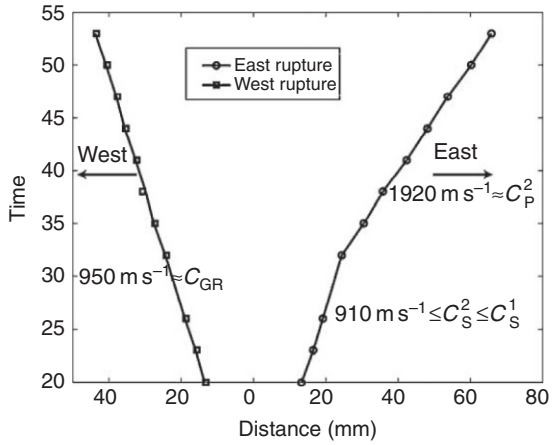


Figure 16 Rupture length plot of an experiment for $\alpha = 25°$, $P = 13$ MPa, and rough surface finish.

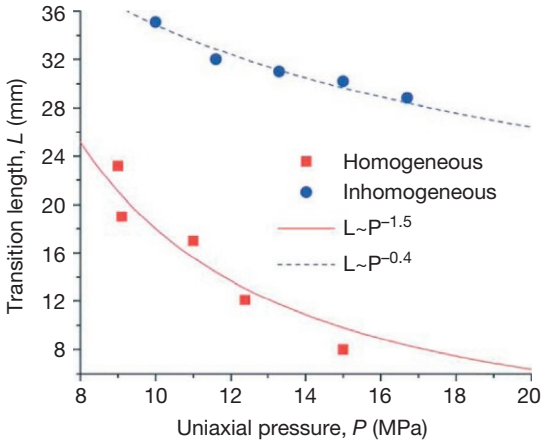


Figure 17 Transition length as a function of pressure P for experiments with $\alpha = 25$. The data and fitting curves for the homogeneous case are included for comparison.

(Xia et al., 2004), while the transition length, L, is also a decreasing function of α and P. Most important to the discussion of this chapter is the observation that the ruptures that propagate to the opposite direction require a certain minimum rupture length before they become supershear. This observation provides a clear intuitive link between super shear growth in the 'opposite' direction and large earthquakes. In contrast, no such transition was observed for 'positively' growing '+GR' ruptures irrespective of α, P, and rupture length. As a result, the experiments do not provide an obvious link between 'positively' growing ruptures and large earthquakes.

In other words, a certain minimum length of rupture growth in the negative direction is required before such a transition can be observed. This suggests that in smaller earthquake events, such as the 2004 Parkfield rupture such a transition may not have had the chance to happen before the negative moving rupture was arrested (for more on Parkfield, see Section 4.08.4.5).

In Section 4.08.3.3, we have discussed the dependence of the transition length L on the uniaxial pressure P. In the homogeneous case there is a very well-defined point for transition, while in the inhomogeneous case the transition point is not always so clear. This difference again is due to the presence of a material contrast. In the homogeneous case, there is an energetically forbidden velocity zone between C_R and C_S (Rosakis, 2002). As a result, the secondary crack is initiated exactly at the shear wave front. In contrast for inhomogeneous fault systems, the forbidden zone no longer exists (Rosakis, 2002) and the subshear crack in the opposite direction accelerates to the supershear speed in a smoother way (Figure 16). Nevertheless, we can still define the transition length where a rapid speed change occurs. The plot of transition length L is in Figure 17. L has a weaker dependence on P ($L \sim P^{-0.4}$) than the homogeneous case ($L \sim P^{-1.5}$). This is expected because for the inhomogeneous case, and for subshear rupture speeds, the coupling between the slip and the normal traction causes a local rupture-tip compression. This compression increases the resistance to slip for rupture growing in the negative direction. On the other hand, the shear traction driving the rupture is constant and hence it takes a longer slip distance for a rupture in bimaterial interface to reach the supershear velocity.

4.08.4.4 Comparison of the Experimental Results to Early Numerical and Theoretical Studies

The experiments by Xia et al. (2004, 2005) described above provide the first full-field and real-time visualization of dynamic frictional rupture events occurring along inhomogeneous interfaces, which feature low-wave-speed mismatch such

that the generalized Rayleigh wave speed can be defined. While it is very difficult to assess whether the ruptures are pulse-like, crack-like, or a mixture of the two, the observations confirm the existence of two distinct self-sustained and constant speed rupture modes. These bare strong similarities to the ones that have been theoretically and numerically predicted over the recent years (Andrews and Ben-Zion, 1997; Ben-Zion, 2001; Ben-Zion and Huang, 2002; Cochard and Rice, 2000; Rice, 2001; Weertman, 1980). In particular, a '+GR' type of rupture mode is always excited instantaneously along the 'positive' direction of sliding. Besides the '+GR' rupture mode, a '$-P_{SLOW}$' mode is observed as long as the rupture propagating in the 'opposite' direction is allowed to grow to sufficiently long distances from the hypocenter. The triggering of the '$-P_{SLOW}$' mode is always preceded by a purely sub-Rayleigh, crack-like rupture whose velocity depends on loading, on geometry, and on the bimaterial characteristics. Therefore, the existence of this preliminary and apparently transient stage is one of the main differences with the early numerical predictions (Andrews and Ben-Zion, 1997; Cochard and Rice, 2000). However, its existence does not contradict early theoretical studies (Adams, 2001; Ranjith and Rice, 2001), which can only predict stable rupture events whose constant velocities relate to the wave speeds of the bimaterial system.

A far more striking difference to some of the early numerical predictions (Andrews and Ben-Zion, 1997; Ben-Zion and Huang, 2002; Cochard and Rice, 2000) is the consistent experimental observation of bilateral slip. In contrast to the experiments, the above numerical predictions seem capable only of exciting one or the other of the two self-sustained rupture modes (Cochard and Rice, 2000), giving rise to purely unilateral rupture events. They also seem to primarily favor the triggering of the '+GR' mode in low-wave-speed mismatch bimaterial systems (Andrews and Ben-Zion, 1997). This kind of preference has led to the labeling of the 'positive' direction as the 'favored' rupture direction and of $+C_{GR}$ as the 'favored' rupture velocity. These numerical results indirectly support the closely related notion of ruptured directionality (McGuire et al., 2002). A notable exception to this rule is provided by the early numerical analysis by Harris and Day (1997), as well as their subsequent work (Andrews and Harris, 2005; Harris and Day, 2005) which consistently reports asymmetric bilateral rupture growth in a variety of low-speed contrast, inhomogeneous fault systems. These results are qualitatively very similar to the experimental observations of cases 1 and 2. In particular, the latter 3-D results by Harris and Day (2005) also clearly report on sub-Rayleigh to supershear transition for the rupture propagating along the negative direction. As briefly discussed by Cochard and Rice (2000) and Xia et al. (2005), the excitation of various modes or their combinations should be related to the details of the numerically or experimentally implemented triggering mechanisms. In an attempt to further reconcile the observed differences between various models and the experiments, Xia et al. (2005) have noted that unstable slip rupture propagation has also been observed (Xia et al., 2004) on homogeneous Homalite/Homalite and Polycarbonate/Polycarbonate interfaces. Such unstable rupture growth would be possible only if there was a substantial reduction of the friction coefficient with slip and/or slip rate, and hence such reduction must be a property of both these model materials when sliding against each other. It is then plausible to assume that a similar reduction of friction coefficient occurs along the Homalite/Polycarbonate interface, and to thus infer that its rupture behavior should not be expected to fully correspond to the idealized models of a dissimilar material interface with constant coefficient of friction (Adams, 1995, 1998; Andrews and Ben-Zion, 1997; Cochard and Rice, 2000; Ranjith and Rice, 2001; Weertman, 1980). Indeed, the stronger the weakening becomes, the reduced frictional resistance is expected to further neutralize the effect of bimaterial contrast on normal shear coupling thus rendering the 'preferable' rupture direction less preferable than originally thought.

The present experiments neither support exclusivity nor show a strong preference for rupture direction. Although they support the idea that frictional ruptures that grow in the positive direction will always do so at a specific constant velocity ($V = +C_{GR}$), they still allow for a significant possibility of self-sustained supershear ruptures growing in the opposite direction. This possibility becomes significant, provided that their transient, sub-Rayleigh precursors grow for a large-enough length and are not arrested prior to transitioning to supershear. The requirement of a critical transition length along the 'opposite' direction provides a link between large earthquakes and the occurrence of self-sustained supershear rupture in the 'opposite' direction. One perhaps can contemplate the existence of a weak statistical preference for positively growing ruptures, since this link to large earthquakes is absent for '+GR' ruptures.

4.08.4.5 The Parkfield Earthquake Discussion in the Context of Experiments and of Recent Numerical Studies

The Parkfield earthquake sequence presents a very interesting case in the context of bimaterial rupture and the issue of the existence, or lack of, a preferable rupture direction. The slip on the San Andreas Fault is right-lateral, and the crust, near Parkfield, on the west side of the fault features faster wave speeds than the east side (Thurber et al., 2003). The three most recent major Parkfield earthquakes ruptured the same section of the San Andreas Fault. The rupture directions of the 1934 and the 1966 events were southeastward (positive direction), whereas the 2004 earthquake ruptured in the opposite direction (negative direction). Following the discussion above, results from the early constant friction coefficient studies in bimaterials would imply that the positive (southeastward) direction of the 1934 and the 1966 events is the preferable direction. According to this notion, the negative (northwestward) direction of the 2004 earthquake would not be favored. The experiments described above (Xia et al., 2005), however, have clearly demonstrated that a rupture in the negative direction can indeed occur at least in model materials. Such a rupture, depending on tectonic loading conditions and on available rupture length, may grow at either sub-Rayleigh or at supershear rupture speeds. Motivated by the most recent 2004 Parkfield event and supported by the Xia et al. (2005) study, Harris and Day (2005) extended their earlier 1997 work to three dimensions and discussed their results in relation to extensive observational evidence collected from the past 70 years of earthquakes at the vicinity of Parkfield. They concluded that

natural earthquake rupture propagation direction is unlikely to be predictable on the basis of bimaterial contrast.

The scientific discussion on the notion of rupture directionality is however far from being over. In a recent paper, Shi and Ben-Zion (2006) have presented an extensive parameter study for various bimaterial contrasts which clearly shows the effect of assumed frictional properties and nucleation conditions in promoting either unilateral (preferred) or bilateral (nonpreferred) rupture in the presence of bimaterial contrast. Their study utilized a classical linear slip-weakening model of friction of the same type used by Harris and Day (1997) and Harris and Day (2005). It shows that for small differences $\mu^s - \mu^d \sim 0.1$ between the static and dynamic coefficient of friction and for nonbiased nucleation events, the bimaterial effect indeed induces ruptures which primarily grow in the positive direction at $+C_{GR}$. These cases are consistent with the constant coefficient friction models. However, when the difference between the two coefficients of friction is taken to be larger than or equal to 0.5 and for 20% shear wave speed contrast, bilateral rupture in both directions becomes possible with speeds equal to $+C_{GR}$ in the positive direction, and depending on the nucleation condition, with either subshear or supershear ($-P_{SLOW}$) speeds in the negative direction. These cases are consistent with the experiments and perhaps with the Parkfield observations. Their results also agree with the above-discussed explanation, originally offered by Xia et al. (2005), regarding the root cause of the differences between the experimental results and the prediction of a preferable direction in the early numerical studies. Indeed the difference of friction coefficient in the experiments is expected to be in the range of 0.4–0.5 (see **Section 4.08.2.2.3** for the homogeneous case) which, within the slip-weakening assumption, would promote bilateral ruptures of various types. Moreover, this recent study does not conclusively settle the issue of rupture direction preference in natural earthquakes where frictional characteristics are generally unknown. Nevertheless it clearly shows that the main arbitrator in whether bimaterial contrast is capable of inducing rupture directionality, through normal to shear coupling, is the frictional law. Clearly, much more work is needed to investigate this phenomenon. The inclusion of realistic frictional laws including strong velocity weakening or enhanced rate and state frictional descriptions which regularize the rupture problem could be a first step. Another step would necessitate the numerical and experimental study of a variety of bimaterial contrasts and of more naturally based nucleation conditions. Nucleation conditions are shown by Lu et al. (2005) (homogeneous case) and by Shi and Ben-Zion (2006) and Rubin and Ampuero (2006) (inhomogeneous case) to have a strong influence on subsequent rupture growth. We believe that transition lengths, for example, may be heavily dependent on nucleation conditions. The recent papers by Andrews and Harris (2005) and Ben-Zion (2006), as well as the work of Rubin and Ampuero (2006), are also very important recent references reflecting the on-going discussion on this subject.

4.08.4.6 Discussion of the Historic North Anatolian Fault Earthquake Sequence in View of the Experimental Results

The 1999, M7.4, Izmit earthquake in Turkey is perhaps a prime example of a recent large earthquake event for which both modes of self-sustained rupture may have been simultaneously present, as is the case in our experiments. The event featured right-lateral slip and bilateral rupture of a rather straight strike-slip segment of the North Anatolian Fault shown in **Figure 18**. As reported by Bouchon et al. (2001), the westward propagating side of the rupture grew with a velocity close to the Rayleigh wave speed, while the eastward moving rupture grew at a supershear velocity that was slightly above the $\sqrt{2}C_S$ times the shear wave speed of crustal rock. The visible difference in the nature of the seismographs from stations ARC and SKR situated almost equidistantly to the hypocenter (star) support their conclusion (see bottom of **Figure 18**). Since the laboratory ruptures of this chapter are intentionally oriented similarly to the Izmit event (lower-wave-speed material to the south as

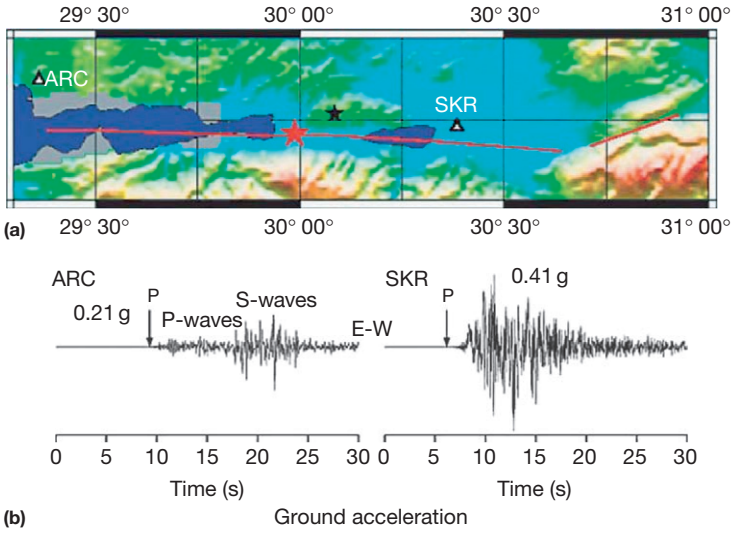


Figure 18 (a) Map view indicating the hypocenter of 1999 Izmit earthquake and two seismological stations ARC and SKR; (b) Seismograms obtained at ARC and SKR. Reproduced from Bouchon M, Bouin MP, Karabulut H, Toksoz MN, Dietrich M, and Rosakis AJ (2001) How fast is rupture during an earthquake? New insights from the 1999 Turkey earthquakes. *Geophysical Research Letters* 28(14): 2723–2726.

indicated by Le Pichon et al., 2003), a direct comparison with the case described in **Figure 14** becomes possible and it reveals some striking similarities between Bouchon et al. (2001) interpretation of the earthquake and the experiment. In addition to featuring right lateral slip and asymmetric bilateral rupture, this experimental case (case-2) featured a subshear westward rupture propagating at $+C_{GR}$. To the east however, the rupture propagated at a velocity slightly lower than C_P^2, which also happens to be equal to 1.6 C_S^1 for the particular bimaterial contrast of the experiments. If one interprets the Izmit event as occurring in an inhomogeneous fault with the lower-wave-speed material being situated at the southern side of the fault (as is in the experiment), the field observations and the experimental measurements of both rupture directions and velocities are very consistent. Moreover, when the bimaterial contrast is low enough, the differences between C_{GR} and the average of the two Rayleigh wave speeds, $\left(C_R^1 + C_R^2\right)/2$, as well as the difference between an inferred rupture speed of 1.6 C_S^1 and $\sqrt{2}\left(C_S^1 + C_S^2\right)/2$ would be small enough not to be discriminated by the inversion process, even if the fault geology was completely known. In this respect, the agreement with experiment is as good as it can ever be expected. In addition, viewing the fault as inhomogeneous can explain the choice of direction for both the sub-Rayleigh and the supershear branches, respectively. This choice of rupture direction is consistent with both the present experiments and with the theories reviewed in the introduction.

The 1999 Düzce earthquake can also be interpreted through a similar line of argument used for Izmit. The Düzce rupture also featured right lateral slip (as did all events that occurred in the North Anatolian Fault between 1939 and 1999) and it extended the Izmit rupture zone 40 km eastward (negative direction) through asymmetric bilateral slip (Bouchon et al., 2001). Thus, similar to the Izmit earthquake, numerical modeling by Bouchon et al. (2001) indicates both sub-Rayleigh westward and supershear eastward rupture fronts. As a result the direct comparison with case-2 described in **Figure 14** provides an explanation for the two rupture directions and respective velocities, similar to the one given for Izmit. This explanation is of course plausible only if one assumes, once again, that the material to the south of the North Anatolian Fault (at its western end) is the lower-wave-speed solid. Strong evidence supporting this assumption has recently been presented by Le Pichon et al. (2003).

By using similar arguments to the ones used for Izmit and Düzce, one can perhaps attempt to provide a unified rationalization of the seemingly random rupture directions and rupture velocities of the interrelated series of earthquakes that occurred since 1939 along the North Anatolian Fault and ended in 1999 with the Izmit and Düzce events. The following argument requires the assumption that, in average and along its entire length, the North Anatolian Fault features the same type of bimaterial in-homogeneity as the one that has been summarized for Izmit and Düzce. However, rather limited evidence supporting such an assumption is currently available (Zor et al., 2006). If in some average sense this is true, one would expect that the slight majority (60%) of the large ($M \geq 6.8$) earthquake events (i.e., (1939-M7.9), (1942-M6.9), (1944-M7.5), (1951-M6.8), (1957-M6.8), (1967-M7.0)) which featured westward growing ruptures, were probably of the '+GR' type. In other words, this assumption implies that they were classical sub-Rayleigh ruptures that moved with velocity equal to $+C_{GR}$ in the 'positive' direction. The remaining ruptures of the series were 'special' in the sense that they featured dominant eastward growth. As previously detailed, out of the remaining four ruptures of the series, the Izmit and Düzce events were bilateral with a western branch of the '+GR' type (consistent with the others) and an eastward, supershear branch of the '$-P_{SLOW}$' type. The 1943 and 1949 ruptures were purely unidirectional and eastward moving; however, their rupture velocities are not known. If these ruptures are to be consistent with the remaining events in the sequence then they could also have developed as the '$-P_{SLOW}$' type. This possibility is more likely for the 1943 event that featured over 250 km of growth length. As estimated in **Section 4.08.2.2.3**, this length is much larger than the critical length required for transition to supershear. By observing that the 1943 and the two 1999 (Izmit and Düzce) events were of a higher magnitude than most of the other events of the complete series, as reported by Stein et al. (1997), further supports the assertion that at least three out of four 'special' events featured partial or total supershear growth along the 'opposite' (eastward) direction. The basic support for this assertion is provided by the experimentally established link between large earthquakes and supershear ruptures growing in the 'opposite' direction, and is consistent with the direct evidence of supershear from the two most recent 'special' events of 1999.

Finally, it should be noted that if earthquakes of lesser magnitude (in the range between M6.4 and 6.8) are also included in the discussion, the North Anatolian Fault series will feature a weak preference for western propagation. This is not very surprising given the above discussed link between large earthquakes and self-sustained supershear along the opposite direction, a link that does not exist for 'positive' (westward growing) '+GR' ruptures. Indeed, in addition to the actual number of ruptures that grew to the east or west, what is of importance here is the actual growth lengths. The results reported by Stein et al. (1997) show that the total length of westward growth is only slightly higher than that of eastward growth. This is again consistent with experiments which show that the self-sustained '+GR' mode is always and instantaneously present after nucleation. In contrast, the self-sustained '$-P_{SLOW}$' mode is often preceded by an unstable subshear phase. For smaller earthquakes this unstable phase may never transition to supershear and instead it may be arrested. This in turn would result in a total eastward rupture length, which is slightly shorter than the total western rupture length of the earthquake series.

4.08.5 Observing Crack-Like, Pulse-Like, Wrinkle-Like, and Mixed Rupture Modes in the Laboratory

The duration of slip at a point on an interface (or fault) in comparison to the duration of the rupture of the entire fault is a central issue to the modeling of earthquake rupture. There are two widely accepted approaches to the description of dynamic sliding (Rice, 2001). The most classic approach uses elastodynamic shear rupture (or crack) models in which the surfaces

behind the leading edge of the rupture (rupture tip) continuously slide and interact through contact and friction. More recently, pulse-like models in which sliding occurs over a relatively small region have been introduced (Heaton, 1990). In these models, the sliding is confined to a finite length which is propagating at the rupture speed and is followed by interfacial locking.

In faults separating identical materials (homogeneous systems) the crack-like mode of rupture has often (but not exclusively) been generated in many numerical simulations of spontaneous rupture when a rate-independent friction law was implemented (Andrews, 1976, 1985; Das and Aki, 1977; Day, 1982; Harris and Day, 1993; Madariaga, 1976; Ruppert and Yomogida, 1992). It has been pointed out, however, that inversions of seismic data for slip histories, from well-recorded events, indicate that the duration of slip at a point on the fault is one order of magnitude shorter than the total event duration. An attempt to explain this phenomenon has given rise to the concept of a pulse-like rupture mode (Hartzell and Heaton, 1983; Heaton, 1982, 1990; Liu and Helmberger, 1983; Mendoza and Hartzell, 1988, 1989). Some eyewitness accounts have also reported short slip durations during some earthquakes (Pelton et al., 1984; Wallace, 1983).

The concept of a pulse-like rupture went against a widely accepted view of how seismic rupture occurs. Its introduction was followed by various efforts to illuminate the physics leading to this process through analytical and numerical investigations. Different mechanisms for 'self-healing' pulse generation along faults in homogeneous systems have been proposed (Heaton, 1990). One postulation is that if the fault strength is low immediately behind the rupture front and is increased rapidly at a finite distance, then slip might be restricted to a short, narrow propagating area (Brune, 1976). Recent theoretical and numerical investigations show that a strong velocity-weakening friction law model could indeed allow for pulse-like behavior of rupture under certain conditions (Zheng and Rice, 1998). However, simulations utilizing velocity weakening have sometimes resulted in crack-like ruptures and sometimes in 'self-healing' pulses (e.g., Beeler and Tullis, 1996; Ben-Zion and Rice, 1997; Cochard and Madariaga, 1994, 1996; Cochard and Rice, 2000; Coker et al., 2005; Lapusta et al., 2000; Nielsen et al., 2000; Perrin et al., 1995). Friction laws along interfaces between two identical elastic solids have to include laboratory-based rate and state evolution features and they must not exhibit illposedness or paradoxical features (Cochard and Rice, 2000; Ranjith and Rice, 2001). It has been proved that generalized rate and state friction laws are appropriate candidates for homogeneous fault systems (Cochard and Rice, 2000; Ranjith and Rice, 2001; Rice et al., 2001; Zheng and Rice, 1998). Within the frame of rate and state friction laws, the following three requirements have to be fulfilled for rupture to occur as a 'self-healing' pulse (Rice, 2001; Zheng and Rice, 1998). One requirement is that the friction law must include strengthening with time on slipped portions of the interface that are momentarily in stationary contact (Perrin et al., 1995). Another is that the velocity weakening at high slip rates must be much greater than that associated with the weak logarithmic dependence observed in the laboratory during low-velocity sliding experiments. This is often termed 'enhanced' velocity weakening. Lastly, the third requirement is that the overall driving stress has to be lower than a certain value, but high enough to allow for self-sustained pulse propagation (Zheng and Rice, 1998).

While strong velocity weakening is one mechanism which explains the onset of short duration slip-pulses along interfaces (faults) which separate similar materials, it is important to note that other mechanisms exist as well. One involves geometric confinement of the rupture domain by unbreakable regions (barrier model). That is, sliding consists of a number of crack-like ruptures of short duration on a small rupture area that are separated by locked regions (Aki, 1979; Papageorgiou and Aki, 1983a,b). In one implementation of this scenario, a pulse-like rupture behavior was found in a 3-D geometry when the rupture process was confined within a long but narrow region by unbreakable barriers (Day, 1982). It was observed that the rupture starts in a classic crack-like mode and it propagates in all directions. After some time, arresting waves arrive from the boundaries and they effectively relock the fault behind the rupture front, resulting in two slip pulses. Alternatively, a rupture nucleates and propagates bilaterally, but may arrest suddenly at a strong barrier at one end. Following its arrest, the reflected waves from the barrier spread back and heal the rupture surface. The combination of the still-propagating end of the rupture with the healing reflected wave forms a moving pulse-like configuration (Johnson, 1990).

In general, narrow slip pulses can be generated during dynamic sliding along interfaces by strongly velocity-weakening friction on a homogeneous system, by strong fault zone heterogeneities (Rice, 2001) or by variations in normal stress along the rupture interface. All of the above conditions (velocity weakening, heterogeneities, and local normal stress variation) can produce slip pulses with low dynamic stress at the active part of the slip. An extensive discussion of the subject is presented by Nielsen and Madariaga (2003). Most relevant to the discussion here is a set of recent finite element calculations which have been carried out by Coker et al. (2005) for a configuration which is very similar to the experimental setup described in the next section. These simulations have shown that it is possible to generate a great variety of dynamic rupture modes (cracks, pulses, and pulse trains) propagating along an interface characterized by a rate-enhanced, rate- and state-dependent frictional law. The choice of rupture speed and rupture mode clearly depends on the load intensity and rate but it is also very sensitive to the type of constitutive law employed, as well as the values of its parameters.

4.08.5.1 Experimental Investigation of Dynamic Rupture Modes

As discussed in **Section 4.08.2.1**, the majority of the existing experimental studies on sliding concentrate on processes occurring over large timescales (large compared to wave transit within the specimen) and are concerned with developing relationships between time-averaged friction data and various system parameters. Many dynamic friction laws motivated by using various experimental configurations and apparati lack the reproducibility of friction data (survey by Ibrahim (1994)) to be of definite value to the theorists. The results of these experiments are multibranched friction versus slip velocity curves, which even for the same material and the same

experimental configuration depend not only on the properties of the frictional interface but also on the dynamic properties of the apparatus, such as mass, stiffness, and damping. This suggests that the friction data obtained in the course of stick–slip motions do not purely reflect the intrinsic properties of the surfaces in contact but are also greatly affected by several of the dynamic variables involved in each particular experimental setup. Another shortcoming of the experiments described above, in relation to their relevance to real earthquakes, is the achievable range of sliding speeds. Typically, natural earthquakes involve sliding velocities on the order of 1–10 m s^{-1} while these experiments involve sliding in the range of 1 μm s^{-1} to 1 mm s^{-1}. An exception to this rule are the experiments of Prakash and Clifton (1993) and Prakash (1998), who employed a plate-impact pressure-shear friction configuration to investigate the dynamic sliding response of an incoherent interface in the microsecond timescale (sliding speeds on the order of 10 m s^{-1}). The experimental results, deduced from the response to step changes imposed on the normal pressure at the frictional interface, reinforce the importance of including frictional memory in the development of the rate-dependent state variable friction models. However, this setup is essentially 1-D. As such it is not able to provide information about the detailed stress field developed along the length of the interface nor to shed light on the nature (e.g., uniformity or lack of) of the rupture process during dynamic sliding. As in many cases described in Section 4.08.2.1, the implicit assumption crucial to the interpretation of the data is that both stress and sliding processes remain spatially uniform. Another exception is the very recent work by Tullis and his coworkers (Di Toro et al., 2004; Tullis and Goldsby, 2003) who have approached seismic slip rates in experiments involving crustal and quartz rock. They have investigated the phenomenon of flash heating as the main mechanism resulting in severe velocity weakening. They have shown that the phenomenon of flash heating (Rice, 1999) can be experimentally linked to the observed behavior of drastic velocity weakening observed in their experiments. Finally, the scarcity of fast sliding velocity data in the open literature has to be emphasized. We believe that this is an essential data set which could be proved as the ultimate arbitrator for the validation of various dynamic sliding models and friction laws.

Rubinstein et al. (2004) conducted experiments to investigate the onset of dynamic sliding. They considered two blocks of PMMA separated by a rough interface and subjected to vertical static compression and to a gradually increased horizontal driving force. During the sliding initiation, they recorded the relative change of the net local contact area as a function of time by measuring the changes of the light intensity transmitted across the interface. The transmitted light was imaged by a camera. Because of the low framing rates of their recording apparatus, their experimental configuration was particularly suited to the study of 'slow' phenomena, with respect to the characteristic wave speeds of the material. They discovered slow waves propagating at speeds between 40 and 80 m s^{-1}, which is one order of magnitude slower than the Rayleigh wave speed of PMMA. These detachment waves, which are perhaps related to Schalamach waves, because of their low propagation speeds (Schallamach, 1971), are very important since they give rise to a significant amount of slip. There are some doubts, however, to whether the characteristics of these waves are determined only by the properties of the surfaces in contact and by the loading. Because of their low rupture speeds, it is also likely that the dynamic properties of the entire specimen may play an important role in the formation of these waves. If the above is true, then the slow detachment waves, though very important to sliding of finite bodies, may be less relevant to natural earthquakes. Another very interesting question, which also remains to be answered, is whether these 'detachment' waves generate opening at the interface, thus justifying their name.

As it is clear from the experiments presented both in Section 4.08.2.1 and above, most of the experimental methods employed up to this point lack the spatial or/and the temporal resolution to address some important questions on dynamic sliding related to the sliding initiation process, to the duration of sliding (or sliding mode), and to the changes in the stress field distribution along the interface. This is particularly true in cases where sliding speeds on the order of 1 m s^{-1}, or above, are involved. To remedy this situation, the physical plausibility of generating pulse-like and crack-like rupture modes along 'incoherent' (frictional) interfaces in homogeneous systems has been investigated by Lykotrafitis et al. (2006a) in a number of well-controlled experiments. The sliding speeds were on the order of 1 m s^{-1}. This sliding speed is of the same order of magnitude as the one expected in most 'large' natural earthquake rupture events. Dynamic photoelasticity combined with a new laser interferometry-based technique has provided direct physical evidence of the rupture mode type, the exact point of rupture initiation, the sliding velocity history at a point on the interface, and the rupture propagation speed. A summary of these experiments and a discussion of the conditions leading to various modes of ruptures are given in the following sections.

4.08.5.1.1 Specimen configuration and loading

Two Homalite-100 plates, subjected to a uniform compressive stress, were frictionally held along the interface. The top plate was also subjected to an asymmetric impact shear loading (Figure 19). Each of the plates was 139.7 mm long, 76.2 mm wide, and 9.525 mm thick. The average roughness of the surfaces in contact was approximately $R_a = 400$ nm. All the experiments were executed at the same external confining stress of $P = 10$ MPa, applied by a calibrated press. The asymmetric impact loading was imposed via a cylindrical steel projectile of diameter 25.4 mm and length 50.8 mm, fired using a gas gun. The impact speeds ranged from 9 to 72 m s^{-1}. A steel buffer 73 mm high, 25.4 mm long, and 9.525 mm thick was attached to the impact side of the upper plate to prevent shattering by direct impact and to induce a more or less planar loading wave.

4.08.5.1.2 Using particle velocimetry to measure slip and opening histories

The time evolution of the dynamic stress field in the bulk was recorded by high-speed digital photography in conjunction with dynamic photoelasticity (see Section 4.08.2.1). In addition, a new technique based on laser interferometry was adapted to locally measure the horizontal and vertical components of the in-plane particle velocity at various points above and below the sliding interface, thus allowing for the measurement of slip and

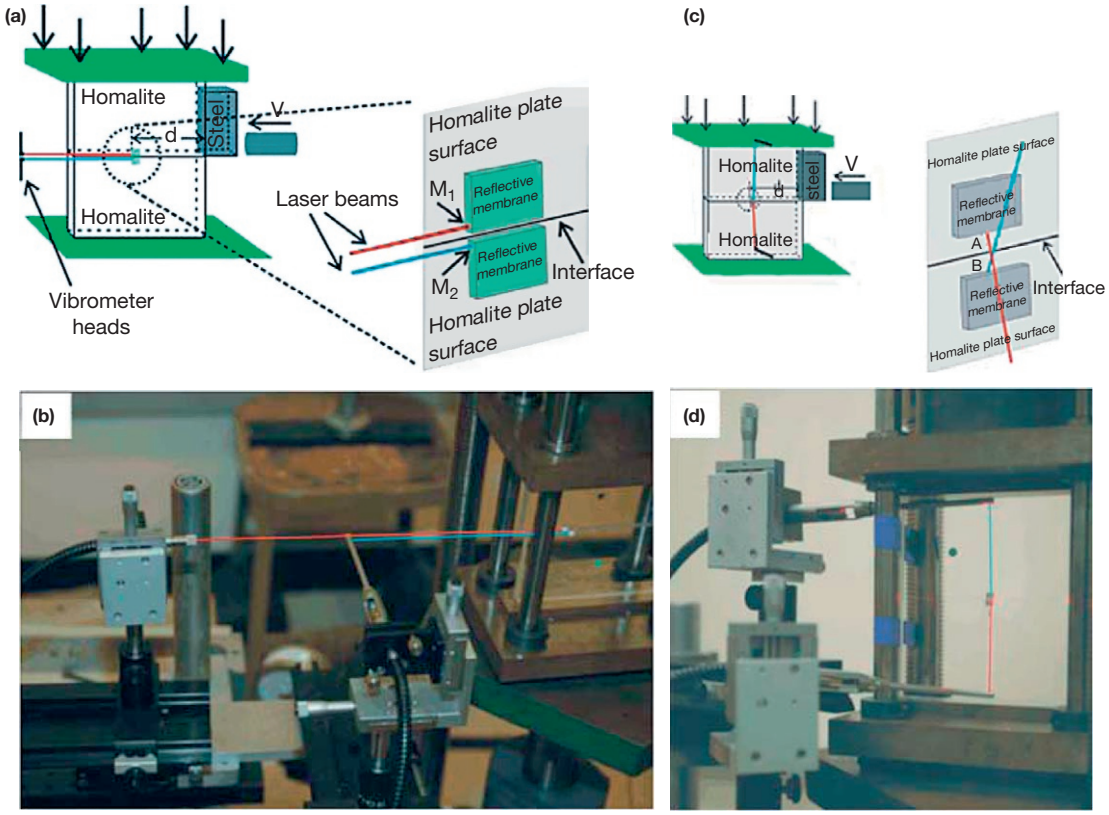


Figure 19 (a) Schematic illustration of the experimental configuration for the sliding velocity measurement. The area inside the dotted circle is shown magnified. Points M_1 and M_2 were at the same horizontal distance d from the impact side of the Homalite plate. (b) Photography of the actual setup displaying the arrangement of the velocimeters' heads for the measurement of the sliding velocity. (c) Schematic illustration of the experimental configuration for the relative vertical velocity measurement. The area inside the dotted line is shown magnified. Points M_1 and M_2 are at the same horizontal distance from the impact side of the Homalite plate. (d) Photography of the actual setup displaying the arrangement of the velocimeters' heads.

opening velocity histories. The combination of the full-field technique of photoelasticity with the local technique of velocimetry is proved to be a very powerful tool in the study of dynamic sliding.

The configuration employed in the measurement of the slip velocity is as follows. A pair of independent fiber-optic velocimeters was used to measure the horizontal particle velocities at two adjacent points M_1 and M_2 across the interface. A schematic illustration of the experimental setup is shown in Figure 19(a), whereas a photograph of the setup is shown in Figure 19(b). The vertical distance of each point from the interface was less than 250 μm. Both points had the same horizontal distance from the impact side of the Homalite plates. By subtracting the velocity of the point below the interface (M_2) from that of the point above the interface (M_1), the relative horizontal velocity history was obtained. The velocimeter consists of a modified Mach–Zehnder interferometer and a velocity decoder. The decoder was set to a full-range scale of $\pm 10\ \mathrm{m\ s^{-1}}$ with a maximum frequency of 1.5 MHz and a maximum acceleration of 10^7 g. The beam spot size was approximately 70 μm, whereas the error of the velocity measurements was 1%. The technique and the corresponding experimental setup are presented in detail in Lykotrafitis et al. (2006b).

Following the same strategy, a pair of independent velocimeters was employed to measure the vertical in-plane components of the velocities of two adjacent points M_1 and M_2, located across the interface. The relative vertical velocity was obtained by adding algebraically the corresponding velocities. The arrangement of the velocimeters is shown schematically in Figure 19(c), whereas a photograph of the actual setup is shown in Figure 19(d). These two types of measurements enabled them to record various modes of rupture propagating along the frictional interface.

4.08.5.2 Visualizing Pulse-Like and Crack-Like Ruptures in the Laboratory

The specimen was subjected to a uniform confining pressure of 10 MPa, whereas the impact speed was $V = 19\ \mathrm{m\ s^{-1}}$. An instantaneous isochromatic fringe pattern is shown in Figure 20(a). An eye-like fringe structure is observed traveling, from right to left, behind the longitudinal wave front. The rupture tip A followed this fringe structure at a supershear speed of 1.36 C_S. Consequently, two Mach lines forming a shear Mach cone emanated from the rupture tip. The tip can be located by tracing the Mach lines to the interface. The rupture-tip speed was found to be constant.

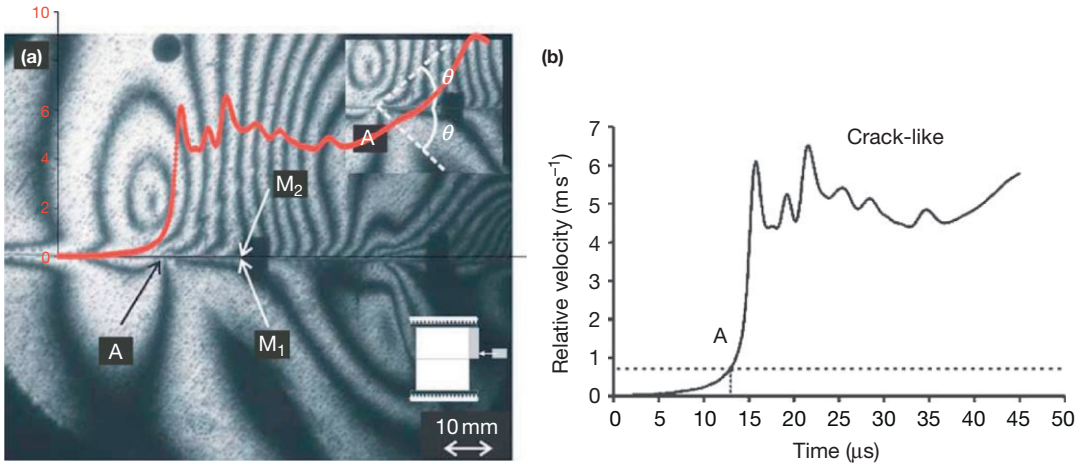


Figure 20 (a) Isochromatic fringe pattern generated during an experiment for which the impact speed was 19 m s^{-1} and the external compression was 10 MPa. The rupture tip is at the fringe concentration point A. The insets highlight the location of the Mach lines emanating from the rupture tip and the specimen configuration. The slip profile is superimposed (red line). (b) Relative velocity history of points M_1 and M_2 located at a distance of 70 mm from the impact side of the Homalite plates. The rupture commenced when the rupture tip A reached the velocity measurement position.

The time evolution of the horizontal relative velocity of two adjacent points M_2 and M_1, belonging to the upper and lower plate, respectively, at a horizontal distance of 70 mm from the impact side of the Homalite plate, is displayed in Figure 20(b). If the entire time history of photoelastic frames is taken into consideration in relation to the velocimeter time record, then this record can be superimposed on the photographs (see red line trace in Figure 20(a)) by converting time to spatial distance.

When the longitudinal wave front arrived at the measurement positions M_1 and M_2 of Figure 20(a), where the pair of the interferometers was pointed, the velocities of both points started to increase. However, the relative horizontal velocity was zero for the next few microseconds and it remained very low for a time interval of approximately 13 μs. A numerical integration of the relative velocity with respect to time from 0 to 13 μs resulted in an accumulated relative horizontal displacement of 2 μm between points M_1 and M_2, which can easily be identified to elastic shear deformation prior to sliding. Indeed, the actual sliding started at approximately 13 μs when the rupture tip (point A as identified by the photoelastic image) arrived at the measurement position and the relative velocity increased sharply. The correlation between the two measurements has allowed us to thus establish an estimate of elastic displacement (∼2 μm) preceding the sliding event for a confining pressure of 10 MPa. After it reached its maximum value of approximately 6 m s^{-1}, the relative velocity decreased and then it fluctuated, but never dropped below 4 m s^{-1} during the recording time. As is evident from Figure 20(b), the sliding was continuous and thus we can safely say that rupture occurred in a classic crack-like mode.

As was already noted, the speed of the rupture tip was substantially higher than the shear wave speed of Homalite, and therefore the situation is similar to intersonic shear rupture propagation observed to occur along both 'coherent' (Coker and Rosakis, 2001; Needleman, 1999; Rosakis et al., 1999, 2000; Samudrala et al., 2002a,b) and 'incoherent' (Xia et al., 2004) interfaces separating identical monolithic solids. In the former experiments, the 'coherent' interfaces were bonded and they featured intrinsic strength and toughness in the absence of confining pressure. Unlike the present study, the resulting modes were always crack-like and the rupture speeds were not constant throughout the event.

In contrast to these early shear crack growth experiments, the present work involves incoherent or frictional interfaces and static far-field compressive loading. Here, the frictional resistance to sliding depends on the normal stress through the friction law. The normal stress, however, is a superposition of the static externally imposed pressure and a dynamic (inertial) compression generated by the impact loading as follows. The P-wave produced by the impact loading creates a primarily horizontal compressive stress in the upper plate close to the interface, and due to the Poisson effect it also creates compression in the direction vertical to the interface. As the sliding proceeds, the vertical stress to the rupture interface changes, and thus the friction changes as well. The change in friction, however, affects the evolution of sliding. Thus, we infer that sliding is dependent on impact loading not only through the horizontal compression, which is the driving force for sliding, but also through the vertical compression which affects the resistance to sliding. Because of that dependence, essential changes in the rupture process are to be expected as the impact speed is decreased.

In order to investigate the above line of reasoning, the impact speed was reduced to 15 m s^{-1}. Figure 21(a) shows an instantaneous isochromatic fringe pattern obtained at the above impact speed. The corresponding relative horizontal velocity history, measured at 30 mm from the impact side of the Homalite plate, is shown in Figure 21(b). By combining the horizontal relative velocity measurement with the recorded photoelastic frames, we can identify two rupture tips A_1 and A_2, which are fringe concentration points and are propagating along the rupture interface at speeds of 1.09 C_S and 0.98 C_S, respectively. The deformation at the velocity measurement position remains elastic until approximately 18 μs when the rupture tip A_1 arrived there and sliding commenced. As in the

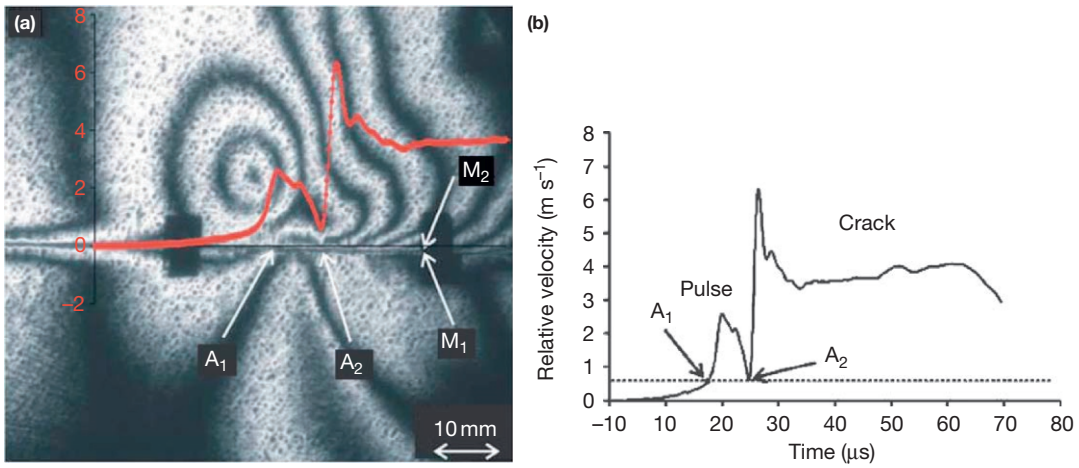


Figure 21 (a) Isochromatic fringe pattern generated during an experiment for which the impact speed was 15 m s^{-1} and the external compression was 10 MPa. The fringe concentration points A_1 and A_2 are the rupture tip and the rear edge, respectively, of the pulse-like rupture mode. The crack-like mode initiated at A_2 right after the pulse. The slip profile is superimposed (red line). (b) Relative velocity history of points M_1 and M_2 located at a distance of 30 mm from the impact side of the Homalite plates. The rupture commenced when the rupture tip A_1 reached the velocity measurement position and a pulse A_1A_2 was formed. The crack-like rupture mode initiated at A_2 right behind the second pulse.

previous case, the commencement of slip corresponds to an accumulated relative horizontal displacement of 2 μm. Subsequently, the horizontal relative velocity increased rapidly from 0.6 m s^{-1} to a local maximum of 2.5 m s^{-1}. After 5 μs, the relative velocity decreased abruptly back to 0.6 m s^{-1} at point A_2 (Figure 21(b)). The slip ceased since the relative velocity was very low, allowing surface asperities to re-establish contact and be deformed elastically. The above observations show that the stable fringe structure (A_1A_2) represents a 'self-healing' slip pulse of approximately 7 μs in duration. Directly after the pulse, the relative velocity increased rapidly again to 6.4 m s^{-1} and retained its large value of approximately 4 m s^{-1} for a long period of time, of about 40 μs. This suggests that the initial rupture of the pulse-like mode was immediately followed by a second rupture of the crack-like mode. Thus, as it has been anticipated, the experimental results presented up to this point indicate that the rupture process is very sensitive to impact speed. Indeed, as the projectile speed was decreased while keeping the external confining stress constant, the rupture-tip speed was also decreased. Finally, the rupture mode changed from a crack-like mode to a mixed mode where a single 'self-healing' slip pulse was followed by a crack.

By further reducing the impact speed to 10 m s^{-1}, the rupture mode became purely pulse-like. The horizontal relative velocity was measured at a distance of 70 mm from the impact side of the Homalite plate and its evolution over time is shown in Figure 22. The rupture started at A_1 and propagated at sub-Rayleigh speed of 0.76 C_S, whereas after 15 μs the sliding ceased at A_2. The duration of sliding was very short compared to the approximately 100 μs duration of the impact event, and thus we infer that an isolated 'self-healing' pulse was formed. Such a case clearly indicates that a pulse-like mode of rupture can definitely occur under appropriate conditions. Indeed this is the first time that such a dynamic pulse has been clearly seen in a controlled laboratory setting. The fact that 'self-healing' pulses were obtained as the impact speed was decreased below certain point is consistent with the theoretical results of Zheng and Rice (1998). According to their predictions, and as discussed in Section 4.08.3.1, one of the necessary requirements of self-healing pulse formation is that the overall driving stress could be lower than a threshold value.

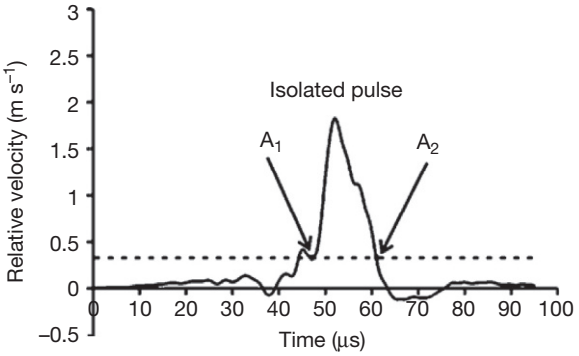


Figure 22 Relative velocity history of points M_1 and M_2 located at a distance of 70 mm from the impact side of the Homalite plates. The rupture commenced when the rupture tip A_1 reached the velocity measurement position and an isolated pulse A_1A_2 was formed.

The combined use of classic dynamic photoelasticity with velocimetry allowed us to fully characterize the frictional sliding process. However, the maximum particle velocity which can be measured by the velocimeter is 10 m s^{-1}. In order to comply with the above limit, low impact speeds were only used up to this point. For higher impact speeds, only dynamic photoelasticity in conjunction with high-speed photography was used as diagnostic tool.

In Figure 23, an instantaneous isochromatic fringe pattern is shown at a specific time for the same compressive load of 10 MPa as in the previous experiments and at much higher impact speed of 42 m s^{-1}. The rupture propagated intersonically at approximately 1948 m s^{-1} = 1.56 C_S = 0.75 C_P giving rise to a shear Mach cone originating from the rupture tip B_1. In addition, a second Mach line which originated from point B_2

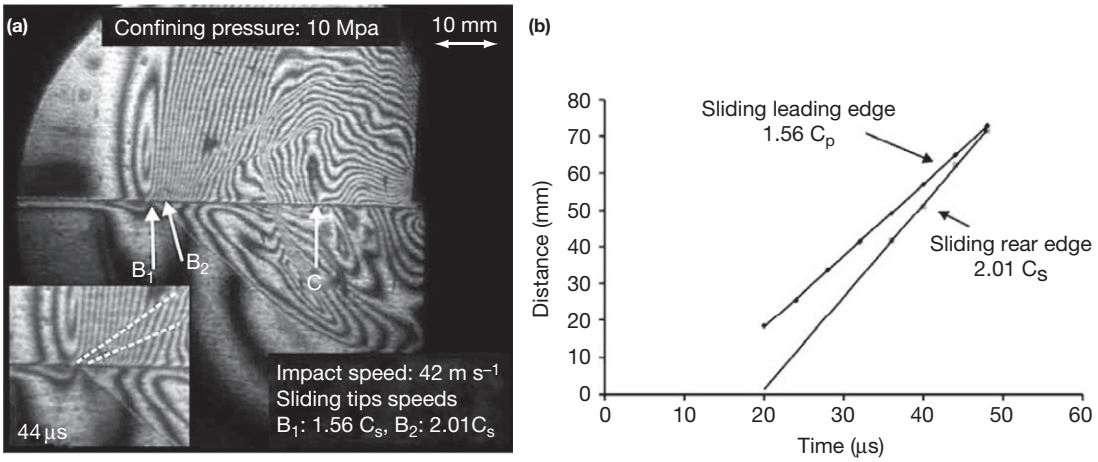


Figure 23 (a) Isochromatic fringes of sliding propagation at different time instances. Two sliding tips (B_1 and B_2) were propagating at different intersonic speeds. In the insets the dual and single Mach lines are marked. (b) Positions of the sliding leading edge and the second sliding tip as functions of time.

and was nonparallel to the first one was observed behind the rupture point (Figure 23(a)). The Mach line was at a shallower slope corresponding to a supershear (approximately sonic) propagation speed of 2514 m $s^{-1} = 2.01\ C_S = 0.97\ C_P$. A more detailed view of the isochromatic fringe pattern in the neighbor of the rupture tip is shown in the inset of Figure 23(a). The two shock waves are highlighted by dotted lines. Nonparallel shock lines imply a highly transient and unstable rupture process. Indeed, study of the entire set of the captured photoelastic pictures shows that the tip B_1 of the second Mach line gradually approached the end B_2 of the first Mach line. Finally, these two points merged as the second point caught up with the first point. The sliding continued at the lower speed and thus only one Mach line was observed in the next recorded frames (not shown here). Figure 23(b) shows the position histories of the first and second sliding tips for the case above. It is evident that the second sliding tip moved faster than the first sliding tip and at approximately 50 μs the pair of points coalesced. The existence of two Mach lines means that behind the onset of sliding (point B_1), there was a second point on the interface (point B_2), where the sliding speed again changed rapidly. Then, one could conjecture that the initial sliding which started at point B_1 stopped after a while and new sliding started at point B_2. In this way, an unstable sliding pulse was formed between the points B_1 and B_2 followed by a crack-like sliding which started at point B_2 in a similar manner to that shown in Figure 21 though at higher sliding speeds. We finally note that behind the second Mach line a 'wrinkle'-like pulse appeared at point C propagating at a speed of 0.92 C_S which is close to the Rayleigh wave speed of Homalite. We extensively elaborate on the nature of this disturbance in Section 4.08.5.3.

An important comment on the frictional sliding experiments discussed above is that the rupture-tip speeds were constants during the entire observation time. That is a general result and it holds true for all the experiments performed. This is a strong characteristic of frictional sliding and it agrees with theoretical results (Adams, 1998; Ranjith and Rice, 2001; Rice et al., 2001) which predict constant and discrete propagation speeds for all the different disturbances and singularities along homogeneous interfaces subjected to uniform prestress.

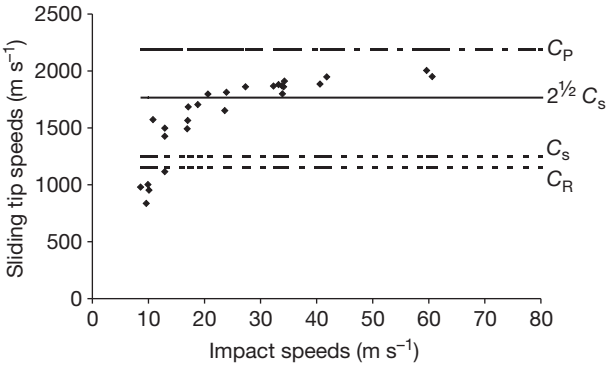


Figure 24 Variation of the sliding tip speed with the impact speed. The confining stress was 10 MPa.

In the last part of this section, the influence of the impact speed on the propagation speeds of the rupture tip is explored. Figure 24 shows the variation of the propagation speed of the sliding front with the impact speed at a constant uniform confining stress of 10 MPa. The slowest achieved impact speed was 9 m s^{-1}. In this case and in other cases with impact speeds close to 10 m s^{-1} the rupture-tip speed was sub-Rayleigh. At higher impact speeds the rupture-tip speed became supershear. It was observed that for impact speeds in the range of 20 to 40 m s^{-1}, the sliding speed was slightly above $\sqrt{2}C_S$. This is a special rupture speed (also discussed in Sections 4.08.3.1, 4.08.4.3, and 4.08.4.5) and it has been shown that it separates regions of unstable and stable intersonic shear crack growth (Samudrala et al., 2002a,b). When the impact speed increased, the rupture-tip speed increased toward the plane stress P-wave speed. It is worth mentioning that no sliding speed was observed in the interval between the Rayleigh wave speed and the shear wave speed of Homalite-100. This experimental observation agrees with theoretical predictions on steady-state

shear crack propagation which exclude this speed interval based on energetic arguments (Broberg, 1999; Freund, 1990; Rosakis, 2002). It should also be noted that the experimental rupture-tip speed measurements presented here feature, as is proven below, high-enough resolution to obtain propagation speeds in the interval between C_R and C_S, if such speeds existed. It is finally noted that experiments performed at lower compression show that the speed of the rupture tip was influenced by the change in the static confining stress and it was supershear even at the lowest achieved impact speed.

4.08.5.3 Wrinkle-Like Opening Pulses Along Interfaces in Homogeneous Systems

The possibility of generating wrinkle-like sliding pulses in incoherent frictionless contact between two dissimilar solids, when separation does not occur, was first investigated by Achenbach and Epstein (1967). These 'smooth contact Stonely waves' (also known as slip waves or generalized Rayleigh waves) are qualitatively similar to those of bonded contact (Stonely waves) and occur for a wider range of material combinations. Comninou and Dundurs (1977) found that self-sustained slip waves with interface separation (detachment waves or wrinkle-like opening slip pulses) can propagate along the interface of two similar or dissimilar solids which are pressed together. The constant propagation speed of these waves was found to be between the Rayleigh wave speed and the shear wave speed of the slowest wave speed material. For interfaces separating identical solids the propagation speed was between C_R and C_S. Weertman (1980) obtained a 2-D self-sustained wrinkle-like slip pulse propagating at the generalized Rayleigh wave speed along a bimaterial interface governed by Coulomb friction when the remote shear stress was less than the frictional strength of the interface. Finite-difference calculations of Andrews and Ben-Zion (1997) show the propagation of wrinkle-like opening pulses along a bimaterial interface governed by Coulomb friction. Particle displacement in a direction perpendicular to the fault is much greater in the slower material than in the faster material, resulting in a separation of the interface during the passage of the slip pulse. Anooshehpoor and Brune (1999) discovered such waves in rubber sliding experiments (using a bimaterial system consisting of two rubber blocks with different wave speeds). The above-mentioned detachment waves are radically different from the Schallamach waves (Schallamach, 1971) which propagate very slowly compared to the wave speeds of the solid.

Lykotrafitis and Rosakis (2006b) observed wrinkle-like opening pulses propagating along the interfaces of Homalite-steel bimaterial structures subjected to impact shear loading. The propagation speeds of these wrinkle-like pulses of finite opening were constants and their values were always between the Rayleigh wave speed and the shear wave speed of Homalite, in accordance with the theoretical prediction of Comninou and Dundurs (1977). The wrinkle-like pulse in the bimaterial case generated a characteristic fringe pattern. Similar photoelastic fringe structures propagating behind the rupture tip along interfaces in homogeneous systems of Homalite were also observed by the same researchers (Lykotrafitis and Rosakis, 2006a). Prompted by this similarity, they decided to investigate whether wrinkle-like pulses were once more responsible for the generation of these fringe structures, in interfaces separating identical solids. Their results are summarized in this section. It is known that a system consisting of two identical half-planes and subjected to compression and to far-field shear loading cannot sustain a wrinkle-like pulse propagating along the interface. In the setup used in these experiments, however, the loading was not strictly shear and there is not any physical reason to exclude the possibility of generating wrinkle-like pulses.

In this section, photoelasticity and velocimetry is used to investigate the physical attainability of such wrinkle-like pulses along frictional interfaces separating identical solids. The equivalent issue for the case of Homalite-steel bimaterial system was investigated by Lykotrafitis and Rosakis (2006b) and Lykotrafitis et al. (2006b). **Figure 25** is used to illustrate the methodology used to identify the wrinkle pulses. The figure features an instantaneous isochromatic fringe pattern obtained at confining stress of 10 MPa and at an impact speed of

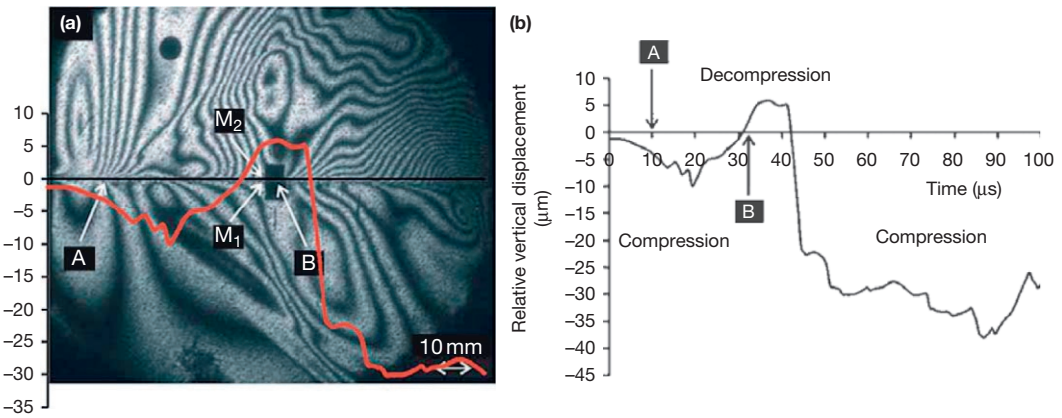


Figure 25 (a) Isochromatic fringe pattern generated during an experiment for which the impact speed was 28 m s^{-1} and the external compression was 10 MPa. The rupture tip is at the fringe concentration point A and the wrinkle-like pulse is at point B. The relative vertical displacement profile is superimposed (red line). (b) Relative vertical displacement history of points M_1 and M_2, located at a distance of 70 mm from the impact side of the Homalite plate. The rupture tip and the wrinkle-like disturbance crossed the velocity measurement position at approximately 10 and 32 μs after triggering, respectively.

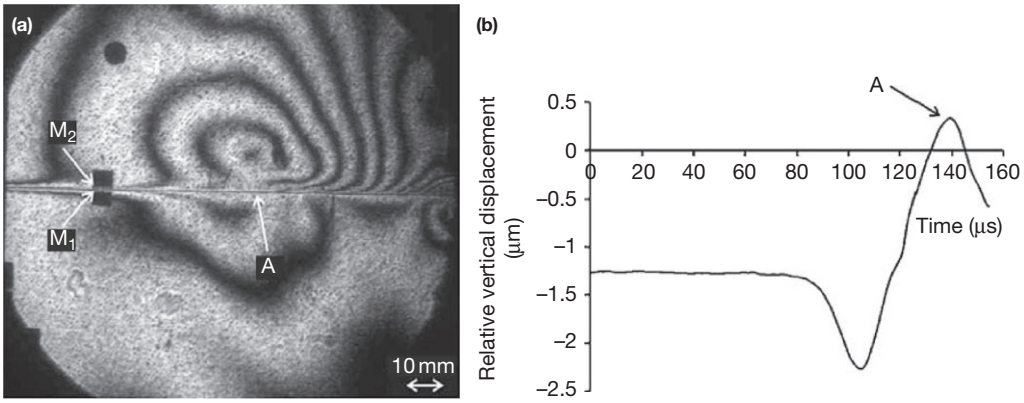


Figure 26 (a) Isochromatic fringe pattern generated during an experiment for which the impact speed was 9 m s^{-1} and the external compression was 10 MPa. The rupture tip is at the fringe concentration point A. No wrinkle-like pulse appeared. (b) Relative vertical displacement history of points M_1 and M_2 located at a distance of 100 mm from the impact side of the Homalite plate.

28 m s^{-1}. Point B is of particular interest. The rupture tip A is shown to propagate at a constant supershear speed of 1.49 C_S. A shear Mach cone is clearly visible in the photoelastic image. A fringe structure located at point B, in Figure 25(a), propagates at a speed of 0.96 C_S. The time evolution of the relative vertical displacement of points M_1 and M_2, which were located 70 mm from the impact side of the Homalite plate, is displayed in Figure 25(b). A simple 1-D calculation shows that the initial static compression of 10 MPa caused a negative relative displacement of approximately 1.3 μm. Negative relative displacement means that the two points approached each other under compression. The P-wave front arrives at about 2.5 μs after the triggering of the oscilloscope. Because of the Poisson effect, the horizontal dynamic (inertial) compression generated a vertical dynamic compression in addition to the static compression. A negative relative displacement was caused by this dynamic compression. The shear rupture point crossed the velocity measurement position at approximately 10 μs (marked A in Figure 25(b)) and it did not cause any visible change to the vertical components of the velocities of points M_1 and M_2. The relative vertical displacement becomes positive at approximately 31 μs. The photoelastic picture, captured at 30 μs, clearly shows that the fringe structure at B, which is the point under investigation, is very close to the measurement position. The study of the entire set of the 16 recorded photoelastic images in combination with the relative vertical displacement history sheds light on the distribution of the dynamic compression along the interface during sliding. The interface was locally under dynamic compression immediately after the arrival of the P-wave front until the arrival of the mentioned fringe formation at B. The entire area from the rupture point A to the location of the fringe structure at B was sliding under compression. The length of the area AB was estimated to be approximately 40 mm, whereas the entire length of the interface was 139.7 mm. The fringe structure at B caused a local opening (~5 μm) at the interface. At approximately 42 μs the relative vertical displacement became again negative and the compression increased abruptly. During the rest of the recording time the interface at the velocity measurement position was sliding under compression. The fringe structure at B can thus be clearly related to a wrinkle-like pulse propagating along the

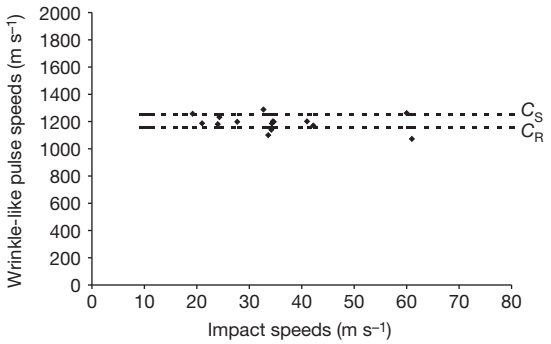


Figure 27 The wrinkle-like pulse speed remained between the Rayleigh wave speed and the shear wave speed of Homalite-100, independent of the impact speed.

interface. As was stated earlier, its propagation speed was very close to 0.96 C_S which lies in the interval between the shear and the Rayleigh wave speeds of the material in agreement with the prediction of Comninou and Dundurs (1977).

Reviewing the results from the entire spectrum of experiments, it is verified that, at a confining stress of 10 MPa, the lowest impact speed which can generate a wrinkle-like pulse in the present setup is approximately 17 m s^{-1}. Figure 26(a) displays an isochromatic fringe pattern obtained at a confining stress of 10 MPa and at an impact speed of 9 m s^{-1}. It is clear that no fringe structure related to a wrinkle-like pulse appears. That is reflected by the relative vertical displacement history shown in Figure 26(b), where the maximum positive value of the displacement was only about 0.3 μm. A comparison of the photoelastic image shown in Figure 26(a) with photoelastic images obtained at the same static compression of 10 MPa and at similar impact speeds during experiments where the horizontal particle velocities measurements were available, shows that the rupture started at point A.

The propagation speeds of wrinkle-like pulses at different impact speeds and at the same confining stress of 10 MPa are shown in Figure 27. The sliding speed was measured to be between the Rayleigh wave and the shear wave speed of Homalite-100. This result, in combination with the

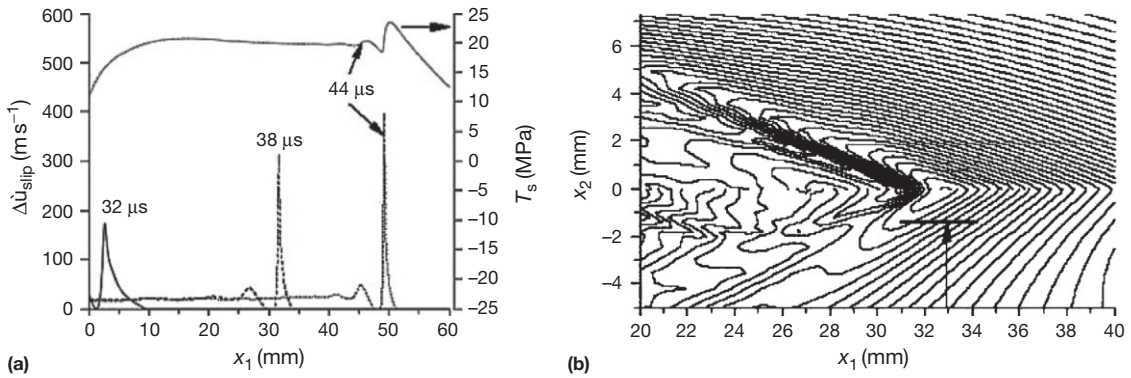


Figure 28 (a) Distributions of the sliding speed $(\Delta\dot{u}_{slip})$ at $t=32$, 38, and 44 μs, and of the shear traction (T_S) at $t=44$ μs along the interface for external compression of 40 MPa and impact speed of 2 m s^{-1}. (b) Contours of maximum shear stress (isochromatic fringe pattern) at $t=38$ μs.

experimental results obtained via velocimetry, favors our conjecture that the interface disturbance was actually a wrinkle-like pulse. As has been already mentioned, the available theoretical and numerical analysis on the subject predicts the same speed range with that identified in Figure 27. It is also noted that both the prestress and the impact speed do not affect the propagation speed of the wrinkle-like pulses provided that such pulses could be generated.

4.08.5.4 Discussion

Finite element calculations of dynamic frictional sliding between deformable bodies have always been a very challenging task for the numerical analysts. As it was recently recognized (see Section 4.08.4.1), the main source of difficulty was the ill-posedness of the corresponding boundary and initial value problem, if a rich-enough friction law was not implemented. Coker et al. (2005) took advantage of the latest advances in the theory of the frictional sliding (Rice, 2001) and by using a rate-enhanced, rate and state friction law of Prakash–Clifton type, they were able to perform stable (grid-size independent) finite element calculations for a configuration very similar to the experiments described above. The use of a rate-enhanced Prakash–Clifton type of law was an essential element for the successful completion of the simulations because of the presence of fast changes of the local compression at the interface, caused by wave propagation, during dynamic sliding. As explained in Section 4.08.4.1, the Prakash–Clifton law is currently considered the only friction law which correctly describes the effect of fast changes in compression on the resulting frictional resistance. We note that for various impact speeds the numerical simulations exhibit a richer behavior than the experimental results and generate not only crack-like, pulse-like, and mixed mode ruptures but also trains of pulses and pulses following a crack-like rupture. A representative result is shown in Figure 28(a), where the distributions of the sliding speed $(\Delta\dot{u}_{slip})$ and the shear traction T_s along the interface at three times ($t=32$, 38, 44 μs) are illustrated, for the case of 40 MPa static compression and of a 2 m s^{-1} impact speed. Contours of maximum shear stress, which correspond to isochromatic fringes in photoelasticity, at $t=38$ μs are shown in Figure 28(b). In this numerical experiment, a mixed mode of rupture, where a pulse was followed by a crack-like rupture, was obtained. This is reminiscent of the experimental case shown in Figure 21. We believe that the similarity between the numerical and the experimental results can be further improved by binding and eventually fine tuning the parameters of the friction law.

The experimental results presented in Section 4.08.5 elucidate the sliding process and provide conclusive evidence of the occurrence of various sliding rupture modes (crack-like, pulse-like, or mixed) propagating dynamically along incoherent interfaces. Of particular interest here is the experimental evidence of the formation of both supershear and sub-Rayleigh sliding pulses of the 'self-healing' type, leading the direct validation of predictions made on the basis of theoretical and numerical models of dynamic shear rupture. These pulses were found to propagate in the absence of interfacial opening. The experiments also provide hints of the dominant physical mechanisms governing the choice of various rupture modes and their evolution. It is finally noted that 'wrinkle-like' pulses which feature finite opening in addition to sliding were discovered propagating along the interface at speeds between C_R and C_S for various loading conditions.

Acknowledgments

The authors gratefully acknowledge the support of NSF (Grant EAR 0207873), the US Department of Energy (Grant DE-FG52-06NA 26209), and the consistent support of the Office of Naval Research (Grant N00014-03-1-0435) and MURI (Grant N000140610730) through Dr. Y.D.S. Rajapakse, Program Manager.

References

Aagaard BT and Heaton TH (2004) Near-source ground motions from simulations of sustained intersonic and supersonic fault ruptures. *Bulletin of the Seismological Society of America* 94: 2064–2078.

Aagaard BT, Heaton TH, and Hall JF (2001) Dynamic earthquake ruptures in the presence of lithostatic normal stresses: Implications for friction models and heat production. *Bulletin of the Seismological Society of America* 91: 1765–1796.

Abraham FF and Gao HJ (2000) How fast can cracks propagate? *Physical Review Letters* 84: 3113–3116.

Achenbach JD and Epstein HI (1967) Dynamic interaction of a layer and a half-space. *Journal of Engineering Mechanics* 5: 27–42.

Adams GG (1995) Self-excited oscillations of two elastic half-spaces sliding with a constant coefficient of friction. *Journal of Applied Mechanics-Transactions of the ASME* 62(4): 867–872.

Adams GG (1998) Steady sliding of two elastic half-spaces with friction reduction due to interface stick–slip. *Journal of Applied Mechanics-Transactions of the ASME* 65(2): 470–475.

Adams GG (2001) An intersonic slip pulse at a frictional interface between dissimilar materials. *Journal of Applied Mechanics-Transactions of the ASME* 68(1): 81–86.

Aki K (1979) Characterization of barriers on an earthquake fault. *Journal of Geophysical Research* 84(6): 140–148.

Andrews DJ (1976) Rupture velocity of plane strain shear cracks. *Journal of Geophysical Research* 81(32): 5679–5687.

Andrews DJ (1985) Dynamic plane-strain shear rupture with a slip-weakening friction law calculated by a boundary integral method. *Bulletin of the Seismological Society of America* 75(1): 1–21.

Andrews DJ and Ben-Zion Y (1997) Wrinkle-like slip pulse on a fault between different materials. *Journal of Geophysical Research: Solid Earth* 102(B1): 553–571.

Andrews DJ and Harris RA (2005) The wrinkle-like slip pulse is not important in earthquake dynamics. *Geophysical Research Letters* 32(23): L23303. http://dx.doi.org/10.1029/2005GL023996.

Anooshehpoor A and Brune JN (1994) Frictional heat-generation and seismic radiation in a foam rubber model of earthquakes. *Pure and Applied Geophysics* 142(3–4): 735–747.

Anooshehpoor A and Brune JN (1999) Wrinkle-like Weertman pulse at the interface between two blocks of foam rubber with different velocities. *Geophysical Research Letters* 26(13): 2025–2028.

Archuleta RJ (1984) A faulting model for the 1979 Imperial-Valley earthquake. *Journal of Geophysical Research* 89(NB6): 4559–4585.

Beeler NM and Tullis TE (1996) Self-healing slip pulses in dynamic rupture models due to velocity-dependent strength. *Bulletin of the Seismological Society of America* 86: 1130–1148.

Ben-Zion Y (2001) Dynamic ruptures in recent models of earthquake faults. *Journal of the Mechanics and Physics of Solids* 49(9): 2209–2244.

Ben-Zion Y (2006) Comment on "Material contrast does not predict earthquake rupture propagation direction" by R. A. Harris and S. M. Day. *Geophysical Research Letters* 33(13): L13310. http://dx.doi.org/10.1029/2005GL025652.

Ben-Zion Y and Huang YQ (2002) Dynamic rupture on an interface between a compliant fault zone layer and a stiffer surrounding solid. *Journal of Geophysical Research: Solid Earth* 107(B2): 2042. http://dx.doi.org/10.1029/2001JB000254.

Ben-Zion Y and Rice JR (1997) Dynamic simulations of slip on a smooth fault in an elastic solid. *Journal of Geophysical Research: Solid Earth* 102(B8): 17771–17784.

Bouchon M, Bouin MP, Karabulut H, Toksoz MN, Dietrich M, and Rosakis AJ (2001) How fast is rupture during an earthquake? New insights from the 1999 Turkey earthquakes. *Geophysical Research Letters* 28(14): 2723–2726.

Bouchon M and Vallee M (2003) Observation of long supershear rupture during the magnitude 8.1 Kunlunshan earthquake. *Science* 301(5634): 824–826.

Bowden FP and Tabor D (1986) *The Friction and Lubrication of Solids*, p. 374. Oxford: Clarenden Press.

Brace WF and Byerlee JD (1966) Stick–slip as a mechanism for earthquakes. *Science* 153(3739): 990–992.

Broberg KB (1999) *Cracks and Fracture*, p. 725. San Diego: Academic Press.

Brune JN (1973) Earthquake modeling by stick–slip along precut surfaces in stressed from rubber. *Bulletin of the Seismological Society of America* 63(6): 2105–2119.

Brune JN (1976) The physics of earthquake strong motion. In: Lomnitz C and Rosenbluth E (eds.) *Seismic Risk and Engineering Decisions*, pp. 141–171. New York: Elsevier.

Brune JN and Anooshehpoor A (1997) Frictional resistance of a fault zone with strong rotors. *Geophysical Research Letters* 24(16): 2071–2074.

Brune JN, Brown S, and Johnson PA (1993) Rupture mechanism and interface separation in foam rubber models of earthquakes – A possible solution to the heat-flow paradox and the paradox of large overthrusts. *Tectonophysics* 218(1–3): 59–67.

Burridge R (1973) Admissible speeds for plane-strain self-similar shear cracks with friction but lacking cohesion. *Geophysical Journal of the Royal Astronomical Society* 35(4): 439–455.

Burridge R, Conn G, and Freund LB (1979) Stability of a rapid mode-I shear crack with finite cohesive traction. *Journal of Geophysical Research* 84(NB5): 2210–2222.

Cochard A and Madariaga R (1994) Dynamic faulting under rate-dependent friction. *Pure and Applied Geophysics* 142(3–4): 419–445.

Cochard A and Madariaga R (1996) Complexity of seismicity due to highly rate-dependent friction. *Journal of Geophysical Research* 105: 25891–25907.

Cochard A and Rice JR (2000) Fault rupture between dissimilar materials: Ill-posedness, regularization, and slip-pulse response. *Journal of Geophysical Research: Solid Earth* 105(B11): 25891–25907.

Coker D, Lykotrafitis G, Needleman A, and Rosakis AJ (2005) Frictional sliding modes along an interface between identical elastic plates subject to shear impact loading. *Journal of the Mechanics and Physics of Solids* 53: 884–922.

Coker D and Rosakis AJ (2001) Experimental observations of intersonic crack growth in asymmetrically loaded unidirectional composite plates. *Philosophical Magazine A: Physics of Condensed Matter Structure Defects and Mechanical Properties* 81(3): 571–595.

Coker D, Rosakis AJ, and Needleman A (2003) Dynamic crack growth along a polymer composite–Homalite interface. *Journal of the Mechanics and Physics of Solids* 51(3): 425–460.

Comninou M and Dundurs J (1977) Elastic interface waves involving separation. *Journal of Applied Mechanics-Transactions of the ASME* 44(2): 222–226.

Dally JW and Riley WF (1991) *Experimental Stress Analysis*, p. 639. New York: McGraw-Hill.

Das S and Aki K (1977) Numerical study of 2-dimensional spontaneous rupture propagation. *Geophysical Journal of the Royal Astronomical Society* 50(3): 643–668.

Day SM (1982) 3-Dimensional simulation of spontaneous rupture – The effect of nonuniform prestress. *Bulletin of the Seismological Society of America* 72(6): 1881–1902.

Di Toro G, Goldsby DL, and Tullis TE (2004) Friction falls toward zero in quartz rock as slip velocity approaching seismic rates. *Nature* 427: 436–439.

Dieterich JH (1972) Time-dependent friction in rocks. *Journal of Geophysical Research* 77(20): 3690–3697.

Dieterich JH (1979) Modeling of rock friction. 1. Experimental results and constitutive equations. *Journal of Geophysical Research* 84(NB5): 2161–2168.

Dieterich JH (1992) Earthquake nucleation on faults with rate-dependent and state-dependent strength. *Tectonophysics* 211(1–4): 115–134.

Dieterich JH and Kilgore BD (1994) Direct observation of frictional contacts—New insights for state-dependent properties. *Pure and Applied Geophysics* 143(1–3): 283–302.

Dunham EM, Favreau P, and Carlson JM (2003) A supershear transition mechanism for cracks. *Science* 299(5612): 1557–1559.

Eberhart-Phillips D and Michael AJ (1998) Seismotectonics of the Loma Prieta, California, region determined from three-dimensional V-p, V-p/V-s, and seismicity. *Journal of Geophysical Research: Solid Earth* 103(B9): 21099–21120.

Ellsworth WL, Çelebi M, Evans JR, Jensen EG, Nyman DJ, and Spudich P (2004) Processing and Modeling of the Pump Station 10 Record from the November 3, 2002, Denali Fault, Alaska earthquake. In: *Proceedings of the Eleventh International Conference of Soil Dynamics and Earthquake Engineering Berkeley, CA, Jan. 7–9*, p. 563.

Freund LB (1990) *Dynamic Fracture Mechanics*. Cambridge: Cambridge University Press.

Fukuyama E and Madariaga R (1998) Rupture dynamics of a planar fault in a 3D elastic medium: Rate- and slip-weakening friction. *Bulletin of the Seismological Society of America* 88(1): 1–17.

Gao HJ, Huang YG, and Abraham FF (2001) Continuum and atomistic studies of intersonic crack propagation. *Journal of the Mechanics and Physics of Solids* 49(9): 2113–2132.

Geubelle PH and Kubair DV (2001) Intersonic crack propagation in homogeneous media under shear-dominated loading: Numerical analysis. *Journal of the Mechanics and Physics of Solids* 49(3): 571–587.

Gu YJ and Wong TF (1994) Development of shear localization in simulated quartz gouge – Effect of cumulative slip and gouge particle-size. *Pure and Applied Geophysics* 143(1–3): 387–423.

Harris RA and Day SM (1993) Dynamics of fault interaction – Parallel strike-slip faults. *Journal of Geophysical Research: Solid Earth* 98(B3): 4461–4472.

Harris RA and Day SM (1997) Effects of a low-velocity zone on a dynamic rupture. *Bulletin of the Seismological Society of America* 87(5): 1267–1280.

Harris RA and Day SM (2005) Material contrast does not predict earthquake rupture propagation direction. *Geophysical Research Letters* 32(23): L23301. http://dx.doi.org/10.1029/2005GL023941.

Hartzell SH and Heaton TH (1983) Inversion of strong ground motion and teleseismic waveform data for the fault rupture history of the 1979 Imperial Valley, California, earthquake. *Bulletin of the Seismological Society of America* 73: 1553–1583.

Heaton TH (1982) The 1971 San Fernando earthquake; a double event? *Bulletin of the Seismological Society of America* 72: 2037–2062.

Heaton TH (1990) Evidence for and implications of self-healing pulses of slip in earthquake rupture. *Physics of the Earth and Planetary Interiors* 64(1): 1–20.

Ibrahim RA (1994) Friction-induced vibration, chatter, squeal, and chaos, Part I: Mechanics of contact and friction. *Applied Mechanics Review: ASME* 47: 209–226.

Ida Y (1972) Cohesive force across tip of a longitudinal-shear crack and Griffiths specific surface-energy. *Journal of Geophysical Research* 77(20): 3796–3805.

Ide S and Takeo M (1997) Determination of constitutive relations of fault slip based on seismic wave analysis. *Journal of Geophysical Research: Solid Earth* 102(B12): 27379–27391.

Johnson E (1990) On the initiation of unidirectional slip. *Geophysical Journal International* 101: 125–132.

Johnson TL and Scholz CH (1976) Dynamic properties of stick–slip friction of rock. *Journal of Geophysical Research* 81(5): 881–888.

Kanamori H (1994) Mechanics of earthquakes. *Annual Review of Earth and Planetary Sciences* 22: 207–237.

Kubair DV, Geubelle PH, and Huang YY (2002) Intersonic crack propagation in homogeneous media under shear-dominated loading: Theoretical analysis. *Journal of the Mechanics and Physics of Solids* 50(8): 1547–1564.

Kubair DV, Geubelle PH, and Huang YGY (2003) Analysis of a rate-dependent cohesive model for dynamic crack propagation. *Engineering Fracture Mechanics* 70(5): 685–704.

Lambros J and Rosakis AJ (1995) Dynamic decohesion of bimaterials – Experimental-observations and failure criteria. *International Journal of Solids and Structures* 32(17–18): 2677–2702.

Lapusta N and Rice JR (2003) Nucleation and early seismic propagation of small and large events in a crustal earthquake model. *Journal of Geophysical Research: Solid Earth* 108(B4): 2205. http://dx.doi.org/10.1029/2001JB000793.

Lapusta N, Rice JR, Ben-Zion Y, and Zheng G (2000) Elastodynamic analysis for slow tectonic loading with spontaneous rupture episodes on faults with rate- and state-dependent friction. *Journal of Geophysical Research* 105: 23765–23789.

Le Pichon X, Chamot-Rooke N, Rangin C, and Sengor AMC (2003) The North Anatolian fault in the Sea of Marmara. *Journal of Geophysical Research: Solid Earth* 108(B4): 2170–2179.

Lin AM, Fu BH, Guo JM, et al. (2002) Co-seismic strike-stip and rupture length produced by the 2001 M-s 8.1 Central Kunlun earthquake. *Science* 296(5575): 2015–2017.

Liu HL and Helmberger DV (1983) The near-source ground motion of the 6 August 1979 Coyote Lake, California, earthquake. *Bulletin of the Seismological Society of America* 73: 201–218.

Lu X, Lapusta N, and Rosakis AJ (2005) Testing friction laws by comparing simulation results with experiments of spontaneous dynamic rupture. *Eos, Transactions of the American Geophysical Union Fall Meet. Suppl.* Abstract S32B-07.

Lykotrafitis G and Rosakis AJ (2006a) Sliding along frictionally held incoherent interfaces in homogeneous systems subjected to dynamic shear loading: A photoelastic study. *International Journal of Fracture* 140: 213–233.

Lykotrafitis G and Rosakis AJ (2006b) Dynamic sliding of frictionally held bimaterial interfaces subjected to impact shear loading. *Proceedings of the Royal Society of London Series A* 462: 2997–3026.

Lykotrafitis G, Rosakis AJ, and Ravichandran G (2006a) Self-healing pulse-like shear ruptures in the laboratory. *Science* 313: 1765–1768.

Lykotrafitis G, Rosakis AJ, and Ravichandran G (2006b) Particle velocimetry and photoelasticity applied to the study of dynamic sliding along frictionally-held bimaterial interfaces: Techniques and feasibility. *Experimental Mechanics* 46: 205–216.

Madariaga R (1976) Dynamics of an expanding circular fault. *Bulletin of the Seismological Society of America* 66: 639–666.

Madariaga R and Olsen KB (2000) Criticality of rupture dynamics in 3-D. *Pure and Applied Geophysics* 157(11–12): 1981–2001.

Magistrale H and Sanders C (1995) P-wave image of the Peninsular Ranges Batholith, Southern California. *Geophysical Research Letters* 22(18): 2549–2552.

McGuire JJ, Zhao L, and Jordan TH (2002) Predominance of unilateral rupture for a global catalog of large earthquakes. *Bulletin of the Seismological Society of America* 92(8): 3309–3317.

Mendoza C and Hartzell SH (1988) Inversion for slip distribution using GDSN P waves: North Palm Springs, Borah peak, and Michoakan earthquakes. *Bulletin of the Seismological Society of America* 78: 1092–1111.

Mendoza C and Hartzell SH (1989) Slip distribution of 19 September 1985 Michoacan, Mexico, earthquake: Near-source and teleseismic constraints. *Bulletin of the Seismological Society of America* 79: 655–669.

Needleman A (1999) An analysis of intersonic crack growth under shear loading. *ASME, Journal of Applied Mechanics* 66: 847–857.

Nielsen SB, Carlson JM, and Olsen KB (2000) Influence of friction and fault geometry on earthquake rupture. *Journal of Geophysical Research: Solid Earth* 105(B3): 6069–6088.

Nielsen S and Madariaga R (2003) On the self-healing fracture mode. *Bulletin of the Seismological Society of America* 93: 2375–2388.

Ohnaka M (2003) A constitutive scaling law and a unified comprehension for frictional slip failure, shear fracture of intact rock, and earthquake rupture. *Journal of Geophysical Research: Solid Earth* 108(B2): 2080. http://dx.doi.org/10.1029/2000JB000123.

Ohnaka M and Shen LF (1999) Scaling of the shear rupture process from nucleation to dynamic propagation: Implications of geometric irregularity of the rupturing surfaces. *Journal of Geophysical Research: Solid Earth* 104(B1): 817–844.

Okubo PG and Dieterich JH (1984) Effects of physical fault properties on frictional instabilities produced on simulated faults. *Journal of Geophysical Research* 89(7): 5817–5827.

Olsen KB, Madariaga R, and Archuleta RJ (1997) Three-dimensional dynamic simulation of the 1992 Landers earthquake. *Science* 278(5339): 834–838.

Palmer AC and Rice JR (1973) Growth of slip surfaces in progressive failure of over-consolidated clay. *Proceedings of the Royal Society of London Series A: Mathematical Physical and Engineering Sciences* 332(1591): 527–548.

Papageorgiou A and Aki K (1983a) A specific barrier model for the quantitative description of inhomogeneous faulting and the prediction of strong ground motion. I. Description of the model. *Bulletin of the Seismological Society of America* 73: 693–722.

Papageorgiou A and Aki K (1983b) A specific barrier model for the quantitative description of inhomogeneous faulting and the prediction of strong ground motion. II. Applications of the model. *Bulletin of the Seismological Society of America* 73: 953–978.

Parsons T, Toda S, Stein RS, Barka A, and Dieterich JH (2000) Heightened odds of large earthquakes near Istanbul: An interaction-based probability calculation. *Science* 288(5466): 661–665.

Pelton JR, Meissner CW, and Smith KD (1984) Eyewitness account of normal surface faulting. *Bulletin of the Seismological Society of America* 74: 1083–1089.

Peltzer G, Crampe F, and King G (1999) Evidence of nonlinear elasticity of the crust from the Mw7.6 Manyi (Tibet) earthquake. *Science* 286(5438): 272–276.

Perrin G, Rice JR, and Zheng G (1995) Self-healing slip pulse on a frictional surface. *Journal of the Mechanics and Physics of Solids* 43: 1461–1495.

Prakash V (1998) Frictional response of sliding interfaces subjected to time varying normal pressures. *Journal of Tribology: ASME* 120: 97–102.

Prakash V and Clifton RJ (1993) Time resolved dynamic friction measurements in pressure-shear. In: Ramesh KT (ed.) *Experimental Techniques in the Dynamics of Deformable Solids Applied Mechanics Division*, pp. 33–48. New York: ASME.

Ranjith K and Rice JR (1999) Stability of quasi-static slip in a single degree of freedom elastic system with rate and state dependent friction. *Journal of the Mechanics and Physics of Solids* 47(6): 1207–1218.

Ranjith K and Rice JR (2001) Slip dynamics at an interface between dissimilar materials. *Journal of the Mechanics and Physics of Solids* 49(2): 341–361.

Renardy M (1992) Ill-posedness at the boundary for elastic solids sliding under Coulomb-friction. *Journal of Elasticity* 27(3): 281–287.

Rice JR (1983) Constitutive relations for fault slip and earthquake instabilities. *Pure and Applied Geophysics* 121(3): 443–475.

Rice JR (1999) Flash heating at asperity contacts and rate-dependent friction. *Eos, Transactions of the American Geophysical Union* 80: F681.

Rice JR (2001) New perspectives in crack and fault dynamics. In: Aref H and Phillips JW (eds.) *Mechanics for a New Millennium (Proceedings of the 20th International Congress of Theoretical and Applied Mechanics, 22 August–September 2000)*, pp. 1–23. Chicago: Kluwer.

Rice JR, Lapusta N, and Ranjith K (2001) Rate and state dependent friction and the stability of sliding between elastically deformable solids. *Journal of the Mechanics and Physics of Solids* 49(9): 1865–1898.

Rosakis AJ (2002) Intersonic shear cracks and fault ruptures. *Advances in Physics* 51(4): 1189–1257.

Rosakis AJ, Samudrala O, and Coker D (1999) Cracks faster than the shear wave speed. *Science* 284(5418): 1337–1340.

Rosakis AJ, Samudrala O, and Coker D (2000) Intersonic shear crack growth along weak planes. *Materials Research Innovations* 3: 236–243.

Rosakis AJ, Samudrala O, Singh RP, and Shukla A (1998) Intersonic crack propagation in bimaterial systems. *Journal of the Mechanics and Physics of Solids* 46(10): 1789–1813.

Rubin AM and Ampuero J (2006) Aftershock asymmetry on a bimaterial interface. *Journal of Geophysical Research* 112(B5): 1978–2012.

Rubin AM and Gillard D (2000) Aftershock asymmetry/rupture directivity among central San Andreas fault microearthquakes. *Journal of Geophysical Research: Solid Earth* 105(B8): 19095–19109.

Rubinstein S, Cohen G, and Fineberg J (2004) Detachment fronts and the onset of dynamic friction. *Nature* 430: 1005–1009.

Ruina A (1983) Slip instability and state variable friction laws. *Journal of Geophysical Research* 88(NB12): 359–370.

Ruppert SD and Yomogida KA (1992) Crack-like rupture model for the 19 September 1985 Michoacan, Mexico, earthquake. *Pure and Applied Geophysics* 138: 407–412.

Samudrala O, Huang Y, and Rosakis AJ (2002a) Subsonic and intersonic mode II crack propagation with a rate-dependent cohesive zone. *Journal of the Mechanics and Physics of Solids* 50(6): 1231–1268.

Samudrala O, Huang Y, and Rosakis AJ (2002b) Subsonic and intersonic shear rupture of weak planes with a velocity weakening cohesive zone. *Journal of Geophysical Research: Solid Earth* 107(B8). http://dx.doi.org/10.1029/2001JB000460.

Schallamach A (1971) How does rubber slide? *Wear* 17: 301–312.

Scholz CH (2002) *The Mechanics of Earthquakes and Faulting*, p. 471. Cambridge: Cambridge University Press.

Scholz C, Molnar P, and Johnson T (1972) Detailed studies of frictional sliding of granite and implications for earthquake mechanism. *Journal of Geophysical Research* 77(32): 6392–6406.

Shi ZQ and Ben-Zion Y (2006) Dynamic rupture on a bimaterial interface governed by slip weakening friction. *Geophysical Journal International* 165(2): 469–484.

Spudich P and Cranswick E (1984) Direct observation of rupture propagation during the 1979 Imperial Valley Earthquake using a short baseline accelerometer array. *Bulletin of the Seismological Society of America* 74(6): 2083–2114.

Stein RS, Barka AA, and Dieterich JH (1997) Progressive failure on the North Anatolian fault since 1939 by earthquake stress triggering. *Geophysical Journal International* 128(3): 594–604.

Thurber C, Roecker S, Roberts K, Gold M, Powell L, and Rittger K (2003) Earthquake locations and three-dimensional fault zone structure along the creeping section of the San Andreas fault near Parkfield, CA: Preparing for SAFOD. *Geophysical Research Letters* 30(3): 1112. http://dx.doi.org/10.1029/2002GL016004.

Tullis TE and Goldsby DL (2003) Flash melting of crustal rocks at almost seismic rates. *Eos, Transactions of the American Geophysical Union* 84(46), Fall Meeting Suppl., Abstract S51B-05.

Uenishi K and Rice JR (2003) Universal nucleation length for slip-weakening rupture instability under nonuniform fault loading. *Journal of Geophysical Research: Solid Earth* 108(B1): 2042. http://dx.doi.org/10.1029/2001JB001681.

Uenishi K, Rossmanith HP, and Scheidegger AE (1999) Rayleigh pulse-dynamic triggering of fault slip. *Bulletin of the Seismological Society of America* 89(5): 1296–1312.

Wallace RE (1983) Eyewitness account of surface fault during the earthquake of 28 October 1983 Borah peak, Idaho. *Bulletin of the Seismological Society of America* 74: 1091–1094.

Weertman J (1980) Unstable slippage across a fault that separates elastic media of different elastic-constants. *Journal of Geophysical Research* 85(NB3): 1455–1461.

Wu FT, Thomson KC, and Kuenzler H (1972) Stick–slip propagation velocity and seismic source mechanism. *Bulletin of the Seismological Society of America* 62(6): 1621–1628.

Xia KW, Chalivendra VB, and Rosakis AJ (2006) Observing ideal 'self-similar' crack growth in experiments. *Engineering Fracture Mechanics* 73: 2748–2755.

Xia KW, Rosakis AJ, and Kanamori H (2004) Laboratory earthquakes: The sub-Rayleigh-to-supershear rupture transition. *Science* 303(5665): 1859–1861.

Xia KW, Rosakis AJ, Kanamori H, and Rice JR (2005) Laboratory earthquakes along inhomogeneous faults: Directionality and supershear. *Science* 308(5722): 681–684.

Zheng G and Rice JR (1998) Conditions under which velocity-weakening friction allows a self-healing versus a crack-like mode of rupture. *Bulletin of the Seismological Society of America* 88: 1466–1483.

Zor E, Sandvol E, Gurbuz C, Turkelli N, Seber D, and Barazangi M (2006) The crustal structure of the East Anatolian plateau (Turkey) from receiver functions. *Geophysical Research Letters* 30(24): 8044. http://dx.doi.org/10.1029/2003GL018192.

This page intentionally left blank

4.09 Slip Inversion

S Ide, University of Tokyo, Tokyo, Japan

© 2015 Elsevier B.V. All rights reserved.

4.09.1	**Introduction**	215
4.09.2	**Construction of Slip Inversion Problem**	217
4.09.2.1	Outline	217
4.09.2.2	Data Preparation	217
4.09.2.3	Setting Fault Models	219
4.09.2.3.1	Parameterization of slip distribution	219
4.09.2.3.2	Example of linear expression: multi-time-window method	221
4.09.2.3.3	Example of nonlinear expression	222
4.09.2.4	Calculation of Synthetic Data	222
4.09.3	**Solving Inverse Problem**	224
4.09.3.1	Best Estimate	224
4.09.3.2	Constraints and Regularization	225
4.09.3.3	Comparison in the Frequency Domain	226
4.09.3.4	Back-Projection Method: An Alternative to Slip Inversion	228
4.09.4	**Example of Slip Model**	229
4.09.4.1	The 1999 Chi-Chi Earthquake	229
4.09.4.2	The 2011 Tohoku-Oki Earthquake	231
4.09.5	**Extended Studies Based on Slip Models**	232
4.09.5.1	Characteristics of Slip Models	232
4.09.5.2	Implication of Slip Models for Fault Dynamics	233
4.09.5.3	Dynamic Modeling and Slip Models	234
4.09.5.4	Scaling of Earthquake Heterogeneity	234
4.09.6	**Discussion and Conclusion**	236
Acknowledgments		236
References		236

Nomenclature

$\mathrm{C}j$	jth constraint matrix	m	Model parameter vector
d^0	Data vector	m_j	jth constraint vector
d^e	Model estimate vector	M_o	Seismic moment (Nm)
Dc	Slip-weakening distance	M_w	Moment magnitude
$\mathrm{E_d}$	Data covariance matrix	N_c	The number of constraints
$\mathrm{E_m}$	Parameter covariance matrix	N_d	Length of data vector
E_s	Seismic energy (J)	N_m	Length of parameter vector
G	Kernel matrix	α_1^2	Hyperparameter $\alpha_1^2=\sigma^2/\sigma_1^2$
G_c	Fracture surface energy (J m^{-2})	σ_1^2	Variance of jth constraint
$G_{ip}(\mathrm{x},t;\xi,\tau)$	Green's function	σ^2	Data variance
$L(\mathrm{m}\|\mathrm{d}^0)$	Likelihood function	$\sigma_{ij}(\mathrm{x},t)$	ij component of stress tensor at $(\mathbf{x}, t)$
$L(\sigma^2,\sigma_1^2)$	Marginal likelihood, $L(\sigma^2,\sigma_1^2)=\int L(\mathbf{m},\sigma^2,\sigma_1^2\|\mathbf{d}^0)p(\mathbf{m}\|\sigma_1^2)\mathrm{d}\mathbf{m}$	$\Delta u_j(\xi,\tau)$	Slip distribution in ith direction at (ξ, τ)
		$\gamma(\mathrm{x})$	Free surface energy, energy to make a new surface

4.09.1 Introduction

Earthquake source is dynamic shear ruptures including fracture and frictional slip on fault planes. To further study the physics of earthquake rupture, we have to know what occurred during an earthquake by resolving the spatial and temporal behaviors of the rupture. The information of the detailed rupture process is also useful for realistic simulations of strong ground motion from a complex source. The rupture process is usually represented by fault slip distribution using a parametric model, which is called a slip model, a heterogeneous slip model, or a finite fault model. Since slip models only describe the rupture history without specific reference to underlying rupture physics, they are also referred to as kinematic models. The method to objectively constrain a slip model by data, usually seismic and/or geodetic data, is slip inversion, which is the central

topic of this chapter. The data are usually contaminated with background noise and our knowledge about earthquake locations and underground structure is not sufficient. To overcome these difficulties, various techniques of slip inversion have been developed with many applications during these decades. We review the history, typical formulation, application examples, and extension of slip inversion, paying attention to the limitation of the method.

First, we review the history of the development of slip inversion. Modern seismology began in the 1960s after the establishment of a basic picture of seismic sources as shear slips on fault planes (see **Chapter 4.02**). Revealing the spatial distribution of the slip soon became a major topic of earthquake seismology. The first earthquake whose spatial extent is quantitatively discussed based on seismograms was the 1960 Chilean earthquake (Mw 9.5), the largest event of the twentieth century. The studies of surface waves using long-period seismograms suggest that this earthquake propagated 750–1000 km to the south at a rupture velocity a little slower than the S wave velocity (Benioff et al., 1961; Press et al., 1961).

In the 1960s, macroscopic seismic source models described by a handful of parameters were well studied. The most famous macroscopic model that includes simple rupture propagation is Haskell's model (Aki and Richards, 2002; Haskell, 1964, 1969). In this model, a line dislocation of a constant amount and a constant duration (rise time) propagates on a rectangular plane unidirectionally at a constant velocity. The 1966 Parkfield earthquake (Mw 6.0) was an example that was successfully explained by a propagating dislocation (Aki, 1968). Haskell's model and similar macroscopic models are used to explain observed records of many large earthquakes during the 1960s and 1970s. Based on the results for earthquakes larger than magnitude 6, Kanamori and Anderson (1975) derived scaling relations for macroscopic earthquake ruptures. In these scaling relations, macroscopic parameters such as the fault length, the fault width, and the average slip satisfy a geometric similarity, indicating that the average stress drop is almost constant. This is a standard macroscopic image of earthquake source established in the 1970s.

However, from the very beginning of the seismic observation history, it has also been known that an earthquake is not a simple rupture but consists of several distinct shocks. Quantitative evidence of such complexity was clearly given from the observation of far-field body waves. Far-field records of some large earthquakes contain multiple pulses that are not explained by the effect of the underground structure. These events are called multiple shocks. One early example is the 1976 Guatemala earthquake (Mw 7.5) studied by Kanamori and Stewart (1978) and Kikuchi and Kanamori (1982, 1991). This earthquake consists of ten discrete ruptures that radiated pulse-like body waves. Although not all earthquakes are that complex, some degree of complexity is visible in any large earthquake.

From the late 1970s, seismologists started developing systematic procedures to analyze the complexity of earthquake source. Trifunac (1974) constructed and solved the first slip inversion problem that objectively determines seismic slip on an assumed fault plane for the 1971 San Fernando earthquake. He divided the assumed fault plane into a number of rectangular subfaults and determined the amount of slip on each of the subfaults using the least-squares method. Following the development of the inversion theory in other fields, such as statistics and information engineering, various systematic procedures have been developed to reveal the heterogeneity of seismic source. In 1982, two important papers for inversion methods were published. One is the subevent deconvolution method proposed by Kikuchi and Kanamori (1982). They consider seismic sources as a sequence of spatially distributed point sources and identified the time and moment of these sources iteratively using waveform correlations of far-field body waves. This method and its extended version were applied for many earthquakes (Kikuchi and Kanamori, 1986, 1991, 1994). The other inversion method proposed by Olson and Apsel (1982) determined spatial and temporal slip distribution on a fault plane using successive time windows propagating at a constant velocity. Hartzell and Heaton (1983) also proposed a similar method. These methods are referred to as the multi-time-window method and have been frequently applied to many earthquakes since the 1979 Imperial Valley earthquake (Mw 6.4). We will review this method and its extension more thoroughly in **Section 4.09.2**.

In the history of slip inversion, a small number of large earthquakes played significant roles to provide new kinds of data, to enable development of special treatments, and to increase our knowledge of earthquake physics. Very near-field records of the 1966 Parkfield earthquake (Mw 6.0), at 80 m from the surface fault trace, confirmed the image of seismic source as a propagating dislocation (Aki, 1968). The 1971 San Fernando earthquake (Mw 6.7) provided more than 250 near-field strong-motion records and enabled the first slip inversion from the data at five stations within the source area as described earlier (see **Chapter 4.18**; Trifunac, 1974). The 1979 Imperial Valley earthquake (Mw 6.4) is the first event whose slip models were published by different groups (Archuleta, 1984; Hartzell and Heaton, 1983; Olson and Apsel, 1982). In the 1990s, geodetic measurements using satellites became popular, and these data are introduced into joint inversions as constraints of final slip. The 1992 Landers earthquake (Mw 7.2), the 1994 Northridge earthquake (Mw 6.7), and the 1995 Hyogo-Ken Nanbu (Kobe) earthquake (Mw 6.9) are well-studied events using both seismic and geodetic data. Serious damage by the Kobe earthquake stimulated the improvement of nationwide observational systems in Japan, the dense networks of high-sensitivity and strong-motion seismometers and global positioning system (GPS), which enabled detailed studies of smaller earthquakes. The 1999 Chi-Chi, Taiwan, earthquake (Mw 7.6) is a well-resolved earthquake owing to the dense strong-motion network in Taiwan. As an example of well-resolved earthquake, we compare three slip models of this earthquake in **Section 4.09.4.1**.

After the 2004 Sumatra (Sumatra–Andaman, Mw 9.3) earthquake, we experienced many very large earthquakes, such as the 2005 Sumatra (Nias, Mw 8.6), the 2010 Chile (Maule, Mw 8.8), and the 2011 Tohoku-Oki (Mw 9.0) earthquakes. These mega-earthquakes were also analyzed for slip models, mainly using teleseismic stations worldwide. In addition, for such great earthquakes, a dense seismometer network can be utilized to reveal the whole rupture process in relatively high frequencies by simple stacking of observed waveforms (Ishii et al., 2005; Krüger and Ohrnberger, 2005).

This is the back-projection method, an alternative way to image seismic sources, which we will discuss in Section 4.09.3.4. The comparison between slip models and back-projection images will be shown for the Tohoku-Oki earthquake in Section 4.09.4.2.

In recent years, slip inversion has become a kind of routine analysis. Rapid solutions are published via the Internet from several groups immediately after large earthquakes, usually within one or a few days. Since there are some distributed computer programs of slip inversion, there is a possibility that it is applied without careful investigation. Sometimes, assumptions to get solutions are not explicitly written in research papers. If there are enough data, the relative significance of the assumptions is low, but the data are often limited and the assumptions can be essential for the details of the solution.

In the following sections, we review general treatments and assumptions in slip inversion decomposing the problem into basic elements: preparation of data, model parameterization, and calculation of synthetic data from the model. All of these are essential to construct even a forward problem that just compares data and model prediction. In Section 4.09.2, we introduce various examples for each element. Section 4.09.2.2 reviews the availability of seismic waveform data and several geodetic measurements. Most popular seismic waveform data are local strong-motion records and global broadband records. In Section 4.09.2.3, we summarize the representation of slip distribution by a finite number of model parameters, which are the models to be determined by slip inversion and generally classified into linear or nonlinear types. Synthetic data are calculated using Green's functions, which are the displacements at stations due to an impulsive force and connect data and model parameters. As we will discuss in Section 4.09.2.4, they are theoretically calculated based on the knowledge of underground structures or empirically modified from the records of small events. Section 4.09.3 explains how to solve the problem based on inversion theory. After a short review of an ordinary solution for the least-squares problem in Section 4.09.3.1, Bayesian modeling for underdetermined problems is reviewed with mathematical expressions in Section 4.09.3.2. While most slip inversion problems are solved in the time domain, sometimes, information in the frequency domain is also important, although there have been only a few studies in the frequency domain. We will discuss these studies in Section 4.09.3.3. Section 4.09.3.4 reviews the back-projection method, a recently developed method to reveal the spatiotemporal rupture process. In Section 4.09.4, we show examples of slip inversion by comparing slip models of the Chi-Chi earthquake (Section 4.09.4.1) and the 2011 Tohoku-Oki earthquake (Section 4.09.4.2) published by different groups. Section 4.09.5 reviews various derivative studies based on slip models, finding characteristics, dynamic implication and modeling, and scaling problem. Finally, Section 4.09.6 summarizes this chapter, discussing the current problems and the future prospects of slip inversion.

There have been good review articles about slip inversion. The study of the 1986 North Palm Springs earthquakes by Hartzell (1989) discussed the merits and demerits for different settings, comparing teleseismic and strong-motion data, linear and nonlinear models, and empirical and theoretical Green's functions. The developments till the 1980s are summarized by Kikuchi (1991) and Iwata (1991). Yoshida (1995) reviewed the problem paying special attention to classify the problems from a point source to a spatiotemporal slip distribution. Beroza (1995) described the characteristics of analysis with far-field and near-field data. The review of strong-motion seismology by Anderson (2003) included a list of previous slip inversion studies using strong-motion records. From the viewpoint of earthquake physics, Kanamori (2004) mentioned characteristics of some results of slip inversion. There are also good textbooks and review articles for general inversion problems in geophysics (e.g., Aster et al., 2012; Matsu'ura, 1991; Menke, 2012; Tarantola, 2005).

4.09.2 Construction of Slip Inversion Problem

4.09.2.1 Outline

First, we introduce the mathematical formulation of slip inversion problems using seismic and/or geodetic data. An earthquake rupture is the spatial and temporal distribution of displacement discontinuity $\Delta\mathbf{u}(\mathbf{x}, t)$ across fault planes Σ. We consider only shear slip whose direction is perpendicular to the local normal vector of a fault plane, $\boldsymbol{\nu}(\mathbf{x})$, so that $\Delta\mathbf{u}(\mathbf{x}, t)\cdot\boldsymbol{\nu}(\mathbf{x})=0$ everywhere. In the elastic medium, displacement at $(\mathbf{x}, t)$ due to a slip distribution $\Delta\mathbf{u}(\boldsymbol{\xi}, \tau)$ is written as (eqn [10.1] of Aki and Richards, 2002)

$$u_i(\mathbf{x},t)=\int_{-\infty}^{\infty}\mathrm{d}\tau\iint_{\Sigma}\Delta u_j(\boldsymbol{\xi},\tau)C_{jkpq}G_{ip,q}(\mathbf{x},t;\boldsymbol{\xi},\tau)\nu_k(\boldsymbol{\xi})\,\mathrm{d}\Sigma(\boldsymbol{\xi}) \quad [1]$$

where C_{jkpq} is the elastic constants. G_{ip} is a Green's tensor function, that is, ith component of displacement at $(\mathbf{x}, t)$ due to an impulsive force at $(\boldsymbol{\xi}, \tau)$ in pth direction. The subscript q represents the derivative in ξ_q direction. The summation convention is used for indices j, k, p, and q. Equation [1] is the expression for seismic waves, but it is also applicable to geodetic measurements if we take $t\to\infty$. The constituents of this problem are data u_i, slip model Δu_j, and Green's functions G_{ip}, which we will discuss independently in the following subsections.

4.09.2.2 Data Preparation

Among various data available for slip inversion, seismic waveform data are essential to resolve temporal change of fault slip. Broadband seismograms recorded by global seismic networks are useful to resolve slip distribution for any earthquake larger than M7. Far-field P and S waves in the range of hypocentral angular distances from 30° to 100° are separated from other large phases and interpreted as a source time function with minor modification for the reflection and refraction due to local structure around the source and receiver. Seismic waveform data from regional networks are useful to improve the resolution of the slip distribution for large earthquakes and enable slip inversion even for earthquakes of moderate size. Strong-motion records are especially most important to get a fine image of rupture propagation for large earthquakes. When on-scale records are available at around surface fault traces, the

near-field components can strongly constrain the timing of local rupture propagation (see Chapters 4.15 and 4.18).

Many slip models have been determined using strong-motion records. The resolution and reliability of each model depend mainly on the quantity and quality of strong-motion data. Digital data generally have wider dynamic range than digitized analog data and the accuracy of absolute timing is often critical for the analysis. Strong-motion data are usually limited and only a small number of earthquakes have been observed at many near-field stations. Table 1 summarizes the slip models determined by the relatively large amount of near-field data, together with the models in early studies for the San Fernando earthquake and the Imperial Valley earthquake. The number of strong-motion stations and types of data increased with time and the number of model parameters also increased. Strong-motion data are ground acceleration in most cases but often used as velocity data after integration or as displacement data after double integration.

Recently, we have been able to easily access several data sets of seismic waves. Global data had been available to world researchers before the first slip inversion study. World Wide Standard Seismograph Network (WWSSN; Oliver and Murphy, 1971) constructed in the 1960s was also useful for the study of seismic sources. This network has been gradually replaced by the network of broadband seismometers with digital recording systems at more than 120 worldwide stations (Lee, 2002). Today, we can download these data from Data Manager Center (DMC) (http://www.iris.edu/) of Incorporated Research Institutions for Seismology (IRIS) a few hours after the occurrence of earthquakes. Recently, near-field strong-motion stations have been organized. Websites are maintained by some large regional networks of strong-motion seismometers such as California Strong Motion Instrumentation Program (CSMIP) (http://www.consrv.ca.gov/cgs/smip/) in the United States and K-NET and KiK-net (http://www.kyoshin.bosai.go.jp/) in Japan. The Consortium of Organizations for Strong-Motion Observation Systems (COSMOS) (http://www.strongmotioncenter.org/vdc) provides a virtual server (search engine) for the strong-motion data from stations worldwide.

While new earthquakes provide new data sets for slip inversion, old analog seismograms were digitized and used for slip inversion of old significant earthquakes. It is a time-consuming task to digitize the seismograms from smoked papers and microfilms and apply appropriate corrections to them. Sometimes, the information necessary for the corrections, such as the pendulum period, magnification, and pen arc length, is missing. Nevertheless, owing to recent progress in image processing technique, the study of such old earthquakes is getting popular. For example, the slip models are determined for the 1906 San Francisco earthquake (Mw 7.7; Song et al., 2005; Wald et al., 1993), the 1923 Kanto, Japan, earthquake (Mw 7.9; Wald and Somerville, 1995), the 1944 Tonankai, Japan, earthquake (Mw 7.9; Ichinose et al., 2003; Kikuchi et al., 2003), and the 1948 Fukui earthquake (Mw 6.8; Ichinose et al., 2005; Kikuchi et al., 1999). Although

Table 1 Well-studied earthquakes and slip models

Earthquake	*Date (year/month/day)*	*Mw*	*Reference*	*Near-field data*		*Other data*	*Model*
				Type[a]	*Frequency (Hz)*	*Type*[b]	*Type*[c]
San Fernando, United States	1971/02/09	6.7	Trifunac (1974)	D (5)	0.11–	–	L (22)
Imperial Valley, United States	1979/10/15	6.4	Olson and Apsel (1982)	A (26)	0.1–0.33	–	L (320)
			Hartzell and Heaton (1983)	V (12)	0.1–1.0	FD (19)	L (336)
			Archuleta (1984)	D (18)	−0.5	–	N (840)
Loma Prieta, United States	1989/10/18	6.9	Beroza (1991)	D (20)	0.3–2.0	–	N (1680)
			Steidl et al. (1991)	V (38)	0.05–1.0	–	N (456)
			Wald et al. (1991)	V (16)	0.1–1.0	FV (16)	L (288)
			Liu and Archuleta (2004)	V (16)	0.05–1.0	–	N (480)
Landers, United States	1992/06/28	7.2	Cohee and Beroza (1994)	D (18)	0.018–0.25	–	L (612)
			Wald and Heaton (1994)	D (16)	0.077–0.5	FD (11), G	L (1116)
			Cotton and Campillo (1995)	D (11)	0.05–0.5	–	N (144)
Northridge, United States	1994/01/17	6.7	Hartzell et al. (1996)	V (35)	0.1–1.0	–	N (784)
			Wald et al. (1996)	V (38)	0.1–1.0	FV (13), G	L (1176)
Hyogoken-Nanbu (Kobe), Japan	1995/01/17	6.9	Horikawa et al. (1996)	V (18)	0.05–0.5	G	L (480)
			Sekiguchi et al. (1996)	D (19)	0.1–1.0	–	L (2240)
			Wald (1996)	V (19)	0.05–0.5	FV (13), G	L (1728)
			Yoshida et al. (1996)	D (18)	−1.0	FD (17), G	L (360)
			Ide and Takeo (1997)	D (18)	0.025–0.5	–	L (3344)
Chi-Chi, Taiwan	1999/09/21	7.6	Chi et al. (2001)	V (21)	0.02–0.5	–	L (8320)
			Ma et al. (2001)	D (21)	No filtering	FD(22), G	L (4032)
			Wu et al. (2001)	V (47)	0.016–0.5	G	L (4840)
			Zeng and Chen (2001)	V (15)	0.05–0.25	G	N (960)
			Ji et al. (2003)	V (36)	−1.6	–	N (1620)
Tottori, Japan	2000/10/06	6.7	Iwata and Sekiguchi (2002)	V (12)	0.1–1.0	G	L (1632)
			Semmane et al. (2005)	D (23)	0.1–1.0	G	N (480)

[a]A: acceleration, V: velocity, D: displacement, the number of stations in parenthesis.
[b]FD: far-field displacement, FV: far-field velocity, the number of stations in parenthesis, G: geodetic data.
[c]L: linear presentation, N: nonlinear presentation, the number of parameters in parenthesis.

the resolution and reliability are not comparable to the state-of-the-art results, analyses of these events are important not only for the study of earthquake physics but also for the regional tectonics and the assessment of long-term earthquake potential.

The final slip distribution showing the total slip during an earthquake can be estimated using geodetic data alone (e.g., Yabuki and Matsu'ura, 1992; Ward and Barrientos, 1986). Geodetic data can be combined with seismic data to better constrain the amount of total slip, which is often poorly constrained from seismic data alone. The examples of geodetic data are triangulation (e.g., Yoshida and Koketsu, 1990), leveling (e.g., Horikawa et al., 1996; Yoshida et al., 1996), GPS (e.g., Horikawa et al., 1996; Wald and Heaton, 1994; Wald et al., 1996; Yoshida et al., 1996), and interferometric synthetic aperture radar (InSAR; Delouis et al., 2002; Hernandez et al., 1999; Salichon et al., 2003, 2004). Geodetic data can be used to determine the slip history of extraordinary slow slip events, such as volcanic deformation (e.g., Aoki et al., 1999) and interplate silent earthquakes (Miyazaki et al., 2003, 2006; Yagi et al., 2001). GPS records sampled at high-frequency rate, such as one sample per seconds, were used as displacement seismometers for slip inversion (Miyazaki et al., 2004; Yokota et al., 2009).

Tide gauge records of tsunamis excited by earthquakes, which provide information about the low-frequency behavior of fault slip, are also used to constrain slip distributions (e.g., Fujii et al., 2011; Romano et al., 2012; Satake, 1989, 1993; Tanioka et al., 1996). Since typical tsunamigenic earthquakes occur outside of regional seismic and geodetic networks, tide gauge data are useful to independently constrain the slip distribution. The recent deployment of ocean bottom sensors in Deep-ocean Assessment and Reporting of Tsunamis (DART) system is quite helpful to constrain fault slip from the ocean where little seismic and geodetic observations are available (Fujii et al., 2011).

Before solving an inversion problem, all these data must be properly preprocessed and arranged into a data vector $\mathbf{d}^0$. Analog data require digitization and digital data often need decimation. The seismometer response should be removed by deconvolution to get displacement or velocity data. In this preprocessing, we can estimate the level of noise, apply a band-pass filter to reduce them, and remove characteristic noises if they exist. Problems in the treatment of digital data, such as deconvolution and filtering, are discussed in detail in Scherbaum (1996). As seen in **Table 1**, the type of data (displacement, velocity, or acceleration) is different among studies, which controls the complexity of the slip model. In most inversion studies, data points are assumed to be independent, but this assumption is incorrect, especially for preprocessed time sequence data. Yagi and Fukahata (2008) demonstrated that a careful evaluation of data covariance is essential to objectively estimate a slip model.

When we use different types of data or data from different stations simultaneously, we must determine the relative weights of these data, which is also related to the evaluation of variance/covariance of data sets. The weight should be determined according to the magnitude of noise in each data set. In addition to natural background noises and instrumental noises, we should consider model errors due to incompleteness of our knowledge to calculate synthetic data using assumed models. In the case of seismic data, the largest source of uncertainty comes from the inaccuracy of Green's functions due to uncertain underground structures. The magnitude of such errors is roughly proportional to the amplitude of calculated synthetic data. Therefore, to equalize the weight of all data, we may normalize data by their maximum value for each station. Furthermore, when the spatial distribution of stations is not homogeneous, we should increase the weights for the stations in more sparsely covered areas to regularize the problem (see **Chapters 3.11** and **3.12**).

4.09.2.3 Setting Fault Models

4.09.2.3.1 Parameterization of slip distribution

When we model an earthquake rupture process, we usually assume one or a few rectangular fault planes in the source region. Each rectangular plane has seven parameters: the location of one corner (latitude, longitude, and depth), the size (length and width), and the orientation (strike and dip angles). With two parameters concerning the slip, the slip amount, and the rake angle, these nine parameters represent the macroscopic image of the static fault slip as defined in Aki and Richards (2002). The determination of these macroscopic parameters had been a major research topic in earthquake seismology till the 1980s, and numerous macroscopic static fault models have been determined for many earthquakes.

To discuss the spatial slip distribution, we usually divide the rectangular fault planes into a number of subfaults. If we emphasize the objectivity of slip models, the macroscopic fault parameters, seven for each plane, should be determined in slip inversion. However, the problem that solves heterogeneous slip distribution and fault geometry simultaneously is nonlinear and intractable. Instead, we usually assume these seven parameters based on the results of macroscopic analysis with low resolution or supplemental information such as the aftershock distribution, the surface fault trace, the underground fault images by structural surveys, and the focal mechanism solutions including the centroid moment tensor (CMT) solution that represents the force system acting at the source (Dziewonski and Woodhouse, 1983).

Figure 1 shows the variety of assumed fault models for the Kobe earthquake from the catalog of Hashimoto (2005). All the assumed planes are in the clouds of aftershocks. The difference of location is as large as 5 km in the southern part because there are fewer stations compared to the urbanized northern area. The final slip distributions of all models have common characteristics such as large shallow slip in the southwestern part and rather deep small slip in the northeastern part. Cho and Nakanishi (2000) used three-dimensional (3-D) moment tensor distribution instead of 2-D planes and also found similar characteristics. The fact that their 3-D distribution of seismic moment coincides with the short-term aftershock distribution implies that the data have enough resolution to determine the location of fault planes. When there are stations near the surface offsets, the geometry of faults at surface needs to be determined in detail, as complex curves. For example, researchers assumed a set of complex fault planes for the Chi-Chi earthquake since it is recorded at many stations near fault traces. However, it is not always clear whether the same geometry continuously extends downward, and the

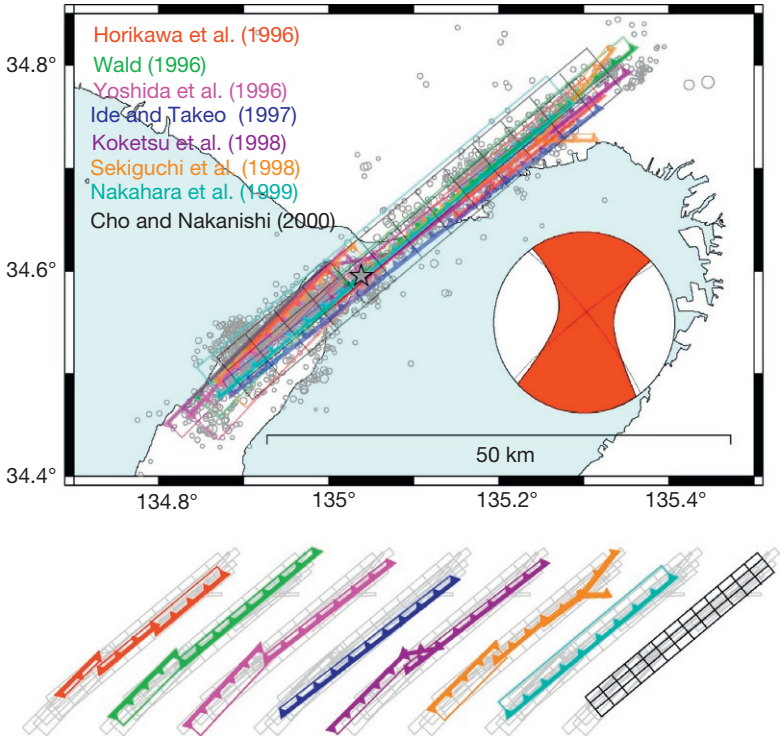


Figure 1 Fault models assumed in the studies of slip inversion by different research groups for the 1995 Kobe earthquake. Except for the model of Cho and Nakanishi (2000), colored bold lines with triangles indicating the direction of dip show uppermost edges and colored thin lines show fault planes projected to the surface. Cho and Nakanishi (2000) used 3-D blocks instead to the fault planes. Star shows the epicenter location determined by Japan Meteorological Agency (JMA). Black circles are aftershock epicenters determined by JMA within 1 week of the mainshock. Harvard CMT solution is shown by a red beach ball. Each of the bottom figures shows one of the fault models by a color and the others by gray lines.

assumption of deep complex fault plane may have unwanted effects on the result.

Once fault planes are assumed, the spatial and temporal slip distribution in the ith direction on each of the fault planes is spatially expanded by J basis functions as

$$\Delta\dot{u}_i(\mathbf{x},t) = \sum_{j=1}^{J} p^j \hat{u}_i^j(t)\phi^j(\mathbf{x})f^j(\mathbf{x},t) \qquad [2]$$

where $\hat{u}_i^j(t)$ is a unit vector representing the slip direction. The direction of slip is always parallel to the fault plane and generally a time-dependent function. $\phi^j(\mathbf{x})$ is a spatial basis function and $f^j(\mathbf{x},t)$ is a slip time function. Both $\phi^j(\mathbf{x})$ and $f^j(\mathbf{x},t)$ are normalized so that the integral of $\phi^j(\mathbf{x})f^j(\mathbf{x},t)$ for space and time is unity. Therefore, the physical dimension of p^j is equal to a volume. A popular rectangular subfault division corresponds to assuming $\phi^j(\mathbf{x})$ as boxcar functions each of which does not overlap with another. Linear or cubic spline functions are other candidates for $\phi^j(\mathbf{x})$. In the case of boxcar functions, p^j represents the seismic potency of the subfault, which is the product of the subfault area and the slip amount. The seismic moment of the subfault is the potency p^j multiplied by the rigidity. Linear or cubic spline functions are often adopted as $\phi^j(\mathbf{x})$. The point subevent representation of Kikuchi and Kanamori (1982, 1991) is essentially equivalent to the use of delta functions as $\phi^j(\mathbf{x})$.

Integrating eqn [2] to $t=\infty$ gives the final slip distribution as

$$\Delta u_i(\mathbf{x}) = \sum_{j=1}^{J} p^j \hat{u}_i^j \phi^j(\mathbf{x}) \qquad [3]$$

where $\hat{u}_i^j$ represents the rake direction of the final slip. We can express

$$\hat{u}_i^j = \mathbf{v}_i^1 \cos\hat{\theta}^j + \mathbf{v}_i^2 \sin\hat{\theta}^j \qquad [4]$$

using two perpendicular unit vectors on the fault plane, $\mathbf{v}^1$ and $\mathbf{v}^2$. $\hat{\theta}j$ is the angle of $\hat{u}^j$ measured from $\mathbf{v}^1$ to $\mathbf{v}^2$. The final slip distribution is determined by potency parameters p^j with a fixed $\hat{u}^j$, or potency parameters $p^{j1} = p^j\cos\hat{\theta}^j$ and $p^{j2} = p^j\sin\hat{\theta}^j$ for two unit vectors. Various geodetic data can constrain these static parameters. Unless we consider $\hat{u}^j$ or $\hat{\theta}^j$ as unknown parameters, the slip distribution is expressed as a linear function of unknown parameters in eqn [3], and we can apply a linear inversion theory that will be described in the next subsection.

To discuss slip history, we have to parameterize the time function of each spatial basis (subfault), $\hat{u}_i^j(t)f^j(\mathbf{x},t)$. Two major ways of parameterization are often used. One is the multi-time-window method in which only slip (potency) amounts are unknowns. The other uses timings of rupture of each subfault as unknowns. Typically, the former is a linear problem with a large number of parameters, while the latter is a nonlinear problem with a small number of parameters. In the following, we explain both methods and discuss the strengths and weaknesses of each method.

4.09.2.3.2 Example of linear expression: multi-time-window method

As reviewed in the introduction, the multi-time-window inversion originates from Olson and Apsel (1982) and Hartzell and Heaton (1983), and it is currently the most popular inversion method of seismic waveform data to estimate a slip model. Figure 2(a) illustrates the multi-time-window method. We expand the time function $\hat{u}_i^j(t)f^j(\mathbf{x},t)$ in eqn [2] using K expansion coefficients q^k and a function $f(t)$ as

$$\hat{u}_i^j(t)f^j(\mathbf{x},t) = \sum_{k=1}^{K} q^k v_i^{jk} f(t-(k-1)\Delta t - t_0(\mathbf{x})) \qquad [5]$$

where the unit vector v_i^{jk} indicates the direction of the slip vector for the jth spatial and kth temporal discrete unit in the spatiotemporal source volume. Before $t=0$, $f(t)=0$, and integration of $f(t)$ up to infinity is unity. Practically, a boxcar function and a linear spline (triangle) function are most frequently used as $f(t)$. Δt is the difference of timing between successive time functions. We force expansion coefficient q^k to satisfy

$$\sum_{k=1}^{K} q^k = 1 \qquad [6]$$

so that the meaning of p^j, potency, is unchanged. $t_0(\mathbf{x})$ is the first time when the fault can slip at location $\mathbf{x}$ and usually assumed that $t_0(\mathbf{x}) = |\mathbf{x}-\mathbf{x}_0|/V_r$, where V_r is the propagation velocity of a hypothetical rupture front from the hypocenter $\mathbf{x}_0$. Therefore, all multi-time windows defined by eqn [5] propagate at this velocity. It should be noted that V_r is different from the true rupture propagation velocity. Rather it is the maximum value of possible rupture velocity from the hypocenter to any point of the fault plane. If the rupture does not start at $t_0(\mathbf{x})$, the actual rupture speed is slower than V_r. When $V_r \to \infty$, the model has no assumption about the hypocenter location, which is useful to check the dependence of the solution on this assumption (e.g., Yoshida, 1992).

Substituting eqn [5] into eqn [2], and decomposing the slip vector into two directions as shown in eqn [4], we have

$$\Delta\dot{u}_i(\mathbf{x},t) = \sum_{j=1}^{J}\sum_{k=1}^{K}\left(p^{jk1}v_i^1 + p^{jk2}v_i^2\right)\phi^j(\mathrm{x}) \\ f(t-(k-1)\Delta t - t_0(\mathbf{x})) \qquad [7]$$

where $p^{jk1} = p^j q^k \cos\theta^{jk}$ and $p^{jk2} = p^j q^k \sin\theta^{jk}$. Namely, p^{jkl} is the seismic potency of the lth slip direction on the jth spatial and kth temporal discrete unit. Equation [7] is the general form of source expression in multi-time-window inversion, where the number of unknowns is $2jk$. The first slip inversion by Trifunac (1974) is categorized as the special form of eqn [7] with $K=1$.

Substitution of eqn [7] into eqn [1] and integration for space and time yield a predictive data vector $\mathbf{d}^e$ as

$$\mathbf{d}^e = \mathbf{G}\mathbf{m} \qquad [8]$$

where $\mathbf{m}$ is a model vector that consists of potency parameters p^{jkl}, and $\mathbf{G}$ includes theoretical waveforms calculated for every discrete unit of spatiotemporal source volume. This linear equation is solved using standard procedures of the linear least-squares inversion (e.g., Menke, 2012) on the assumption that the difference between observation and model prediction follows the normal distribution with zero mean. However, a flexible model requires a large number of model parameters. In recent analysis, the number of parameters is sometimes larger than 3000–4000 (Table 1). Usually, the number of data points is much larger than the number of parameters, but these data points are almost surely dependent and the problem tends to be underdetermined. The effective rank of matrix $\mathbf{G}$ is less than the number of parameters, and the best estimates of model parameters are, even if they were obtained, unstable with quite large standard deviations. Such problems require regularization with various constraints, which will be discussed in Section 4.09.3.2. Another deficiency of the multi-time-window method is that it limits the timing of slip. When the duration of the temporal basis functions is too short and real slip occurs

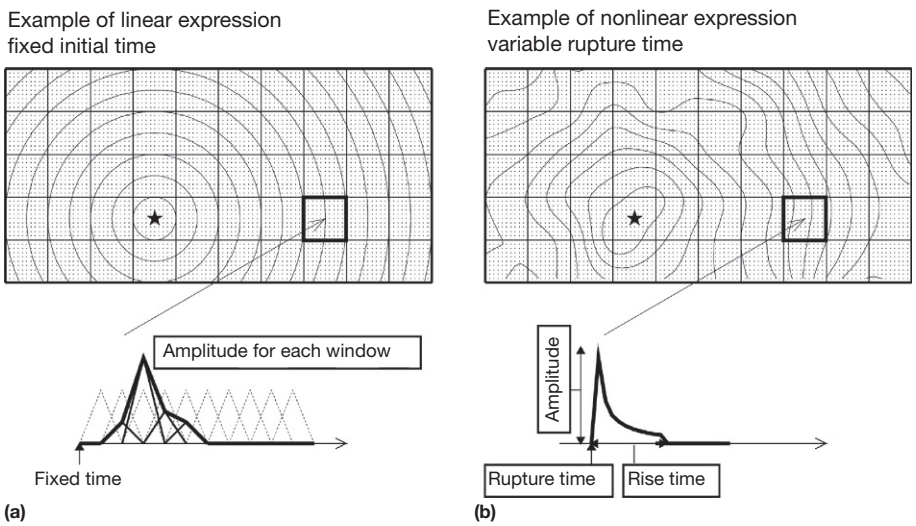


Figure 2 Schematic illustrations showing how to parameterize spatiotemporal slip distribution. The initial time of time function is shown by contours in each of the top figures. (a) Typical linear expression of the multi-time-window analysis. Unknown parameters are amplitude for each window. (b) An example of nonlinear expression, in which rupture time is also an unknown parameter.

longer than that, unexpected errors occur in the solution. This problem was demonstrated by Hartzell and Langer (1993) using the models having different numbers of time windows.

4.09.2.3.3 Example of nonlinear expression

If we assume slip occurs just once at every point on the fault plane and treat the timings of slip as unknown parameters in eqn [2], we can significantly reduce the number of parameters. An example is

$$f^j(\mathbf{x},t) = f\left(\frac{t - t^j - t_0(\mathbf{x})}{r^j}\right) \quad [9]$$

where $f(t)$ is a function which is zero before $t=0$ ($f^j(\mathbf{x},t)=0$ before $t^j+t_0(\mathbf{x})$) and normalized to be unity when integrated with time (**Figure 2(b)**). The parameter r^j is a scaling parameter in time. When $f(t)$ is a boxcar function from $t=0$ to 1, r^j is called the rise time. A boxcar function (e.g., Yoshida, 1986) and a triangle function (e.g., Fukuyama and Irikura, 1986; Takeo, 1987) are frequently used. A merit of nonlinear expression is that we can use a dynamically plausible function as $f(t)$. Beroza and Spudich (1988) assumed

$$f(t) \propto t^{-1/2}[H(t) - H(1-t)] \quad [10]$$

where $H(t)$ is a step (Heaviside) function. This is a truncated version of the slip-rate function $t^{-1/2}$ in the analytic solution of dynamic crack propagation. We can change the direction of slip vector using unknown parameters θ^j as

$$\Delta\dot{u}_i(\mathbf{x},t) = \sum_{j=1}^{J} p^j \nu_i(\theta^j)\phi^j(\mathbf{x}) f\left(\frac{t - t^j - t_0(\mathbf{x})}{r^j}\right) \quad [11]$$

Substituting [11] into [1] and integrating for space and time yields a predictive data vector from the model parameter vector $\mathbf{m}$ consisting of p^j, t^j, r^j, and θ^j:

$$\mathbf{d}^e = \mathbf{G}(\mathbf{m}) \quad [12]$$

If we assume Gaussian errors of model prediction, linear inversion can provide a unique solution, while nonlinear inversion does not eliminate the problem of nonuniqueness. Instead, the model with nonlinear expression can be extended to include more general features of fault slip.

4.09.2.4 Calculation of Synthetic Data

To carry out integration of eqn [1] and obtain predictive data vectors [8] or [12], Green's function must be properly prepared for each point on the fault plane. Unless the spatial basis function $\phi^j(\mathbf{x})$ in eqn [2] is a delta function, we have to calculate Green's functions from every point source located at a small interval to account for rupture directivity effect in each discrete unit (subfault) as shown in **Figure 2**.

Theoretical formulation of elastic wave propagations from a point source in layered structure had been developed mainly during the 1970s and 1980s. For global studies, the method using a propagator matrix (Bouchon, 1976) and generalized ray theory (Langston and Helmberger, 1975) is frequently used to calculate far-field P and S waves with important contributions from local source and site structures, such as depth phases, pP, sP, and sS. For calculation of near-field full waveforms, the method that uses frequency–wave number integration is effective. The reflectivity method (e.g., Fuchs and Müller, 1971; Kind, 1976) and the Kennett–Bouchon method (e.g., Bouchon, 1981; Kennett and Kerry, 1979; Yao and Harkrider, 1983) are well-known methods and several numerical codes have been developed and distributed (e.g., Kohketsu, 1985; Saikia, 1994; Takeo, 1985). A general textbook about seismic wave propagation has been published by Kennett (2001). The combination of ray theory and isochrone integration (Bernard and Madariaga, 1984; Spudich and Frazer, 1984) is also effective in high frequency for near-field stations. For static deformations, the most popular study is Okada (1985) that gives analytic solutions for the surface displacement due to a buried dislocation on a rectangular fault in half-space, with a concise summary of previous works (see **Chapters 1.04** and **1.05**).

The seismic velocity structure profiles determined by explosion surveys are helpful to construct a layered structure. Most studies assume a 1-D structure from the complex result of 2-D or 3-D surveys with minor improvements by trial and error. When the structure difference is large among stations, several 1-D structures are assumed in one study. Waveforms of small events are useful to empirically improve 1-D structure as shown by Ichinose et al. (2003) who compared synthetic waveforms with aftershock records to estimate average structures between the source and the stations.

Though the 1-D structure is thought to be a fair approximation in many cases, it is obvious that there are many problems for which 1-D structure is no longer valid. For example, body waves from submarine earthquakes are followed by large reverberations from water layers, and dip angle of seafloor is critical to simulate these waves (Okamoto and Miyatake, 1989; Wiens, 1987, 1989; Yoshida, 1992). Another example is the existence of thick sediment layers recently revealed by structural surveys in urbanized areas. Using a finite difference method for 3-D structure, Graves and Wald (2001) and Wald and Graves (2001) demonstrated that the use of 3-D Green's functions can increase the resolution of slip inversion. They also emphasized that this improvement is only possible with the accurate underground structure and careful examination of its validity is required for slip inversion. As demonstrated by Tsuboi et al. (2003) for the 2002 Denali Fault earthquake (Mw 7.9), one of the fastest computers can calculate seismic waves up to 5 s in a laterally heterogeneous global Earth model. The introduction of more accurate 3-D Green's functions into slip inversion is promising.

Accurate 3-D structures have been determined only for small areas on Earth. Moreover, small earthquakes radiate high-frequency waves, which cannot be modeled theoretically. An alternative of theoretical Green's function is empirical Green's function (eGf), which is usually the record of a small earthquake. This concept is originally proposed by Hartzell (1978) for strong-motion simulations. Note that the displacement of small earthquake is not a Green's function in the strict sense in eqn [1], where Green's function is defined using an impulsive single force, while an earthquake has a double couple force system with a finite duration. Nevertheless, as long as the source duration of the small event is negligible, the difference of the force system is practically insignificant for slip inversion. Fukuyama and Irikura (1986) applied eGf to the 1983 Japan Sea earthquake (Mw 7.7). The large source area of this event requires two aftershocks each of which represents the north and

the south subfaults, since an eGf is valid only for the limited space. For a large earthquake, which breaks shallow faults, the difference of distance from the free surface is hard to neglect, and a number of aftershocks are required to obtain a complete set of accurate Green's functions, which is almost impossible. Since it is also known that aftershocks are rare in the area where the mainshock slip was large, it is hard to find appropriate small earthquakes for the eGf method. Unfortunately, the waves from this most important part of the fault plane are most difficult to simulate using eGf. Nevertheless, in the study with relatively low resolution especially for small earthquakes, eGf is an effective tool as shown by many studies (e.g., Courboulex et al., 1996; Fletcher and Spudich, 1998; Hellweg and Boatwright, 1999; Ide, 2001; Ide et al., 2011; Mori and Hartzell, 1990; Okada et al., 2001). In these studies, seismic waves at a station can resolve slip distribution mainly in the direction of the ray from the source to the station. To better resolve the slip distribution, a set of various ray directions is required. **Figure 3** is an example of the comparison between waveforms from a large event (the 2003 Miyagi-Oki, Japan, earthquake; Mw 7.1) and its possible eGf event of the similar mechanism. The waveforms of eGf events have short durations and site-specific complexity mainly due to local shallow structure. The mainshock records have some distinct phases and the arrivals of the second phases suggest a northward rupture propagation, which is consistent with the slip models presented by Okada and Hasegawa (2003) and Wu and Takeo (2004).

Misalignment between predicted and observed data can be a significant source of errors though it is not always carefully discussed in research papers. Usually, the arrival times calculated for an assumed structure have some errors, and sometimes, data do not have accurate information of absolute timing. Therefore, we have to arrange predicted and observed data to minimize the difference between the corresponding reference phases. Usually, S waves are dominant in near-field displacement records, but the exact determination of S wave arrival times is difficult because it is masked by P wave coda. It is plausible to estimate the S wave arrival times using the records of small events in an empirical manner. Sometimes, poorly determined timings in near-field records lead to considerably different results of slip inversion. This might be one reason of large discrepancies between slip models for the 1999 İzmit earthquake (Mw 7.4; Bouchon et al., 2002; Delouis et al., 2002; Gülen et al., 2002; Sekiguchi and Iwata, 2002; Yagi and Kikuchi, 2000) as pointed out by Beresnev (2003) and Ide et al. (2005). As mentioned earlier, eGf approximates the shape of Green's functions only crudely except for the hypocenter of the eGf event that occurred. Nevertheless, the probability of misalignments is small since the predicted and observed data can be aligned using absolute timing.

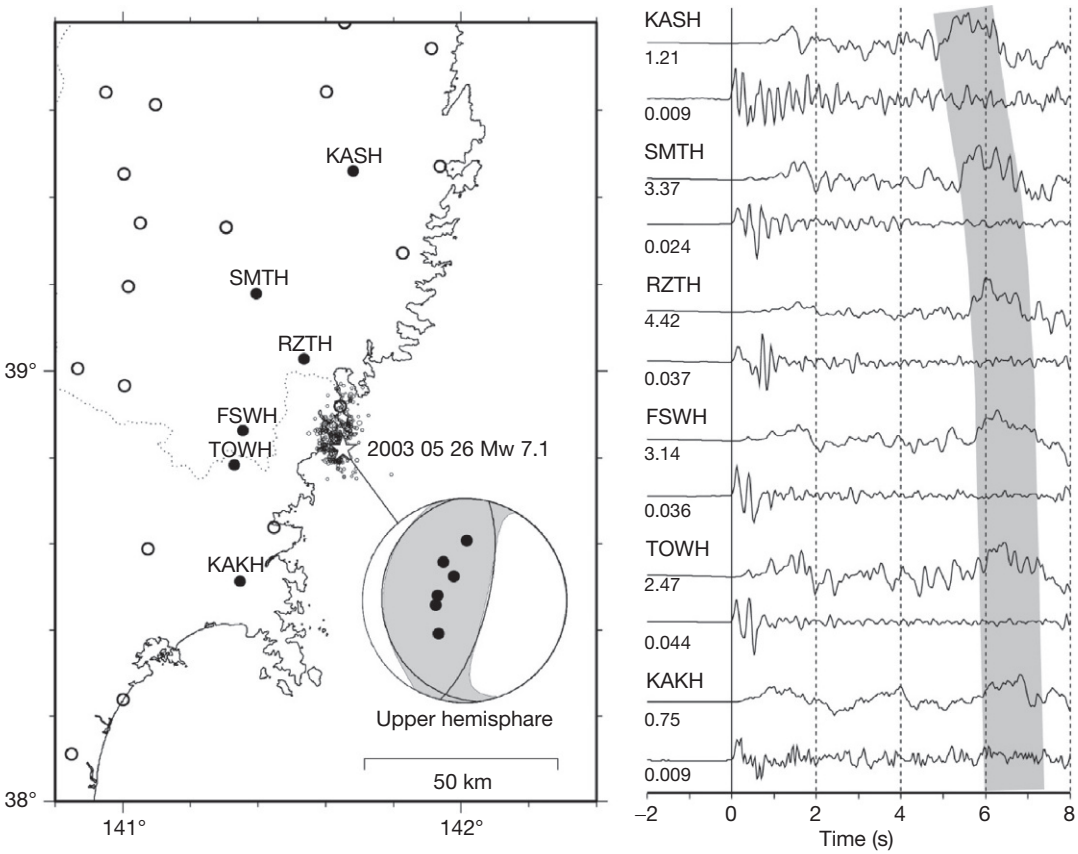


Figure 3 Comparison of waveforms of a large earthquake and its aftershocks, the 2003 Miyagi-Oki, Japan, earthquake. From all available Hi-Net stations (circles), we showed records at six stations (black circles) located in the map and on the focal sphere. For each station, upper and lower traces show the displacement (cm) of the mainshock and the aftershock, respectively. While the records of aftershock are impulsive, the waves of the mainshock last longer. In the shaded area, we can identify the second large phases that indicate rupture propagation northward.

4.09.3 Solving Inverse Problem

4.09.3.1 Best Estimate

From an observed data vector $\mathbf{d}^0$ and a corresponding prediction vector $\mathbf{d}^e$ including a model parameter vector $\mathbf{m}$, we solve the inverse problem to minimize the difference between these vectors, $\mathbf{d}^0-\mathbf{d}^e$, and find the best estimate of the model parameters and their estimation errors. There are good textbooks for such inversion or optimization problems (e.g., Aster et al., 2012; Menke, 2012; Tarantola, 2005). Here, we review basic expressions using the maximum likelihood methods as an introduction to the next subsection.

The first step of inversion problem is to estimate the errors of the data vector. Usually, the errors are assumed to have the Gaussian distribution with zero mean and a covariance matrix. In the next subsection, we review the Bayesian method in which the absolute value of covariance matrix is also estimated from the observed data through marginal likelihood. It is also necessary to estimate relative errors between independent data sets that can have different physical dimensions. For example, we estimate relative magnitude of errors between displacements of two stations, taking into account background noise levels and inaccuracy of the assumed model, that is, the assumption of fault planes and velocity structure. We can assess the effect of model inaccuracy by calculation using slightly different model settings. The relative magnitude of errors between data of different types, for example, seismic waveforms and static displacements, can be estimated similarly.

The maximum likelihood method gives the best estimate of model parameter vector by maximizing a likelihood function $L(\mathbf{m}|\mathbf{d}^0)$. It is formally identical to the probability distribution function of the data $p(\mathbf{d}^0|\mathbf{m})$, but it is thought as a function of $\mathbf{m}$ for the observed data $\mathbf{d}^0$. With the assumption of the Gaussian error, it is

$$L\left(\mathbf{m}\middle|\mathbf{d}^0\right)=\frac{1}{\sqrt{(2\pi\sigma^2)^{N_d}|\mathbf{E}_d|}}\exp\left(-\frac{\left(\mathbf{d}^0-\mathbf{d}^e(\mathbf{m})\right)^t\mathbf{E}_d^{-1}\left(\mathbf{d}^0-\mathbf{d}^e(\mathbf{m})\right)}{2\sigma^2}\right) \quad [13]$$

where we assume that the mean values and covariance matrix of the data as $\mathbf{d}^0$ and $\sigma^2\mathbf{E}_d$, respectively. Superscript t denotes the transpose of a matrix or a vector. N_d is the number of data points. The maximum value of the likelihood function corresponds to the minimum of the L_2 norm in the exponent $(\mathbf{d}^0-\mathbf{d}^e(\mathbf{m}))^t\mathbf{E}_d^{-1}(\mathbf{d}^0-\mathbf{d}^e(\mathbf{m}))$. Therefore, this is the well-known weighted least-squares problem, and the best estimate (maximum likelihood values) of the model parameter, $\mathbf{m}^0$, is given by the least-squares solution.

When the model is linear as in eqn [8], that is, $\mathbf{d}^e(\mathbf{m})=\mathbf{G}\mathbf{m}$, and the matrix $\mathbf{G}^t\mathbf{E}_d^{-1}\mathbf{G}$ is full rank, the solution is

$$\mathbf{m}^0=\left(\mathbf{G}^t\mathbf{E}_d^{-1}\mathbf{G}\right)^{-1}\mathbf{G}^t\mathbf{E}_d^{-1}\mathbf{d}^0 \quad [14]$$

The covariance matrix of the model parameter is written as

$$\mathbf{E}_m=\sigma^2\left(\mathbf{G}^t\mathbf{E}_d^{-1}\mathbf{G}\right)^{-1} \quad [15]$$

However, the solution in the form [14] and [15] is rarely used. One reason is that usually the matrix $\mathbf{G}^t\mathbf{E}_d^{-1}\mathbf{G}$ is singular. The regularization to increase the rank will be discussed in the next subsection. Another reason is that eqn [14] often gives the slips in the negative direction. Although it is still debated whether slip can occur in the negative direction, it is hard to imagine that slip reverses its direction under high shear stress environment in the crust, and so we usually assume that the slip is unidirectional. To prohibit slip in the negative direction, the nonnegative least squares (NNLS; Lawson and Hanson, 1974) are useful. If the solution of NNLS exists, the uniqueness of the solution is verified by the Kuhn–Tucker theorem. A similar positive constraint is also possible with a nonlinear error function that rapidly increases near zero when we solve a nonlinear problem, for which the uniqueness is hard to prove. For example, Yoshida and Koketsu (1990) adopted an error function, $1/\sqrt{X}$, which rapidly increases as slip rate $X\rightarrow 0$.

When the model expression is nonlinear (eqn [12]), it is generally difficult to find the best estimates that minimize $(\mathbf{d}^0-\mathbf{d}^e(\mathbf{m}))^t\mathbf{E}_d^{-1}(\mathbf{d}^0-\mathbf{d}^e(\mathbf{m}))$. We can linearize an observation equation and apply an iterative inversion method such as the Gauss–Newton method and the Levenberg–Marquardt method, in which the solution is derived through successive iterations of linear inversion. The estimate of the parameter vector after ith iteration $\mathbf{m}_i$ gives a prediction vector $\mathbf{d}^e(\mathbf{m}_i)$. Then, the difference between the data vector and the prediction vector is approximated by Taylor series to the first order as

$$\mathbf{d}^0-\mathbf{d}^e(\mathbf{m}_i)\approx\left(\frac{\partial\mathbf{d}^e}{\partial\mathbf{m}}\right)_{\mathbf{m}_i}\delta\mathbf{m}_i=\mathbf{G}'\delta\mathbf{m}_i \quad [16]$$

where $\mathbf{G}'$ is a Jacobian matrix. In the Gauss–Newton method, the solution is derived similarly to eqn [14] and the updated model parameter vector is represented as

$$\begin{aligned}\mathbf{m}_{i+1}&=\mathbf{m}_i+\delta\mathbf{m}_i\\&=\mathbf{m}_i+\left(\mathbf{G}'^t\mathbf{E}_d^{-1}\mathbf{G}'\right)^{-1}\mathbf{G}'^t\mathbf{E}_d^{-1}\left(\mathbf{d}^0-\mathbf{d}^e(\mathbf{m}_i)\right)\end{aligned} \quad [17]$$

This procedure is iterated till the solution converges. However, the convergence does not guarantee that the solution is the best estimate because the solution depends on the initial value and can be just a local least-squares solution. In the case of slip inversion with many unknown parameters, many local least-squares solutions may reduce $L2$ norm similarly, and it is almost impossible to find a unique solution that globally minimizes the $L2$ norm. Jackson and Matsu'ura (1985) showed that prior information for model parameters, as discussed in the following section, can reduce the nonuniqueness.

Rapid improvement of computers has enabled seismologists to solve nonlinear problems by methods including random search iterations such as the Monte Carlo method. As examples to slip inversion in the early period, Hartzell and Liu (1995) and Hartzell et al. (1996) analyzed the Landers earthquake using simulated annealing (SA), and Zeng and Anderson (1996) used a genetic algorithm (GA) for the Northridge earthquake. Ji et al. (2002a) and Liu and Archuleta (2004) recommended the heat-bath algorithm (Rothman, 1986) that is a class of SA method and efficient for slip inversion. The Monte Carlo method also enables the direct estimation of the probability distribution of slip together with the relative weights among different types of data (Fukuda and Johnson, 2010).

Despite the frequent assumption of the Gaussian errors, which is supported by the central limit theorem, it is also

known that real data include many 'outliers' that are far from the prediction and the errors cannot be explained by the Gaussian distribution. One remedy is just to remove these outliers from the data set, but it is also effective to use a long-tailed probability function instead of the Gaussian distribution, which gives more robust solution relatively independent of the outliers. If we assume that the probability function is double exponential (Laplace) distribution, a typical long-tailed function, the likelihood is written as

$$L\left(\mathbf{m}\middle|\mathbf{d}^0\right) = \frac{1}{(2\sigma)^{N_\mathrm{d}}}\exp\left(-\frac{1}{\sigma}\sum_{i=1}^{N_\mathrm{d}}\left|\mathbf{d}_i^0 - \mathbf{d}_i^\mathrm{e}(\mathbf{m})\right|\right) \quad [18]$$

The maximum likelihood solution minimizes the $L1$ norm:

$$\sum_{i=1}^{N_\mathrm{d}}\left|\mathbf{d}_i^0 - \mathbf{d}_i^\mathrm{e}(\mathbf{m})\right| \quad [19]$$

This is called the least absolute values problem, which is solved using an iterative method. There is no dependence on the initial value like general nonlinear problem, and the solution is obtained as a point or a finite hyperplane in the model space (e.g., Tarantola, 2005). The examples of slip inversion using $L1$ norm are found in Das and Kostrov (1990) and Hartzell et al. (1991) who solved the problem by the linear programming method and in Ihmlé (1998) who used SA. The comparison of two solutions based on $L1$ and $L2$ norms is useful to find universal features of the solution that are independent of the norm assumption (Hartzell et al., 1991).

4.09.3.2 Constraints and Regularization

Since the parameterization of slip distribution depends on the specific setup of the model, such as the size and locations of subfaults and the shape of time function, we may use a large number of model parameters to make a flexible model. However, with a large number of model parameters, the solution in the form [14] and [15] cannot be obtained or become unstable with large parameter errors. This is because the effective rank of matrix $\mathbf{G}^\mathrm{t}\mathbf{E}_\mathrm{d}^{-1}\mathbf{G}$ is not full. Although the number of data N_d usually exceeds the number of model parameters N_m, the effective rank becomes smaller because the data are not independent or because a part of the model cannot be resolved by the data. Such a problem requires regularization with some constraints to model parameters. One of the simplest constraints is minimizing the $L2$ norm of the parameter vector. Another popular assumption in slip inversion is that the slip distribution is smooth temporally and spatially in some degree, which is expressed by constraints minimizing the time derivatives and the Laplacian of the slip distribution, respectively. Alternatively, we can fix the total amount of slip or seismic moment at a value suggested from other information (Das and Kostrov, 1990; Ji et al., 2002b), for example, the seismic moment of CMT solution determined using global long-period seismic waves (Dziewonski and Woodhouse, 1983).

The introduction of constraints raises another problem: How strong should these constraints be? A strong smoothing constraint makes the solution smooth, and vice versa. While early studies introduced ad hoc criteria, Bayesian modeling and the information theory provide a way to find an objective solution with a proper weight. In the theory of Bayesian statistics, these constraints are thought to be prior information that provides a prior probability function (pdf) of model parameters with hyperparameters (Jackson, 1979; Jackson and Matsu'ura, 1985). When there is no prior information (constraints), we find the best estimate of model parameter vector maximizing the likelihood function as shown in Section 4.09.3.1. Similarly, with constraints as prior information, we find the solution maximizing a posterior pdf, which is proportional to the product of the likelihood function and the prior pdf. The proper weights of constraints, the best estimates of hyperparameters, are determined by maximizing the expectation of the likelihood for model parameters, a marginal likelihood. The condition is rewritten in a compact form of the minimum Akaike's Bayesian information criterion (ABIC; Akaike, 1980). The ABIC was first applied by Yabuki and Matsu'ura (1992) and many applications of ABIC to slip inversion are found in, for example, Yoshida and Koketsu (1990), Ide and Takeo (1997), and Fukahata et al. (2004). We review the usage following Akaike (1980) and Yabuki and Matsu'ura (1992).

We may use several constraints of different types using constraint matrices $\mathbf{C}_j$ and constraint value vectors $\mathbf{m}_j$ for jth constraint. A constraint is written with errors of a Gaussian distribution with zero mean and a covariance matrix $\sigma_j^2\mathbf{I}$, $\mathbf{e}_j \sim N(0, \sigma_j^2\mathbf{I})$. $\mathbf{C}_j\mathbf{m} = \mathbf{m}_j + N(0, \sigma_j^2\mathbf{I})$. To minimize the $L2$ norm of the parameters, we use a unit matrix for $\mathbf{C}_j$ and a zero vector for $\mathbf{m}_j$. A typical smoothing constraint that minimizes the difference between two successive parameters is

$$\mathbf{C}_j = \begin{pmatrix} 1 & 0 & 0 & \cdots \\ -1 & 1 & 0 & \\ 0 & -1 & 1 & \\ \vdots & & & \ddots \end{pmatrix} \quad [20]$$

and $\mathbf{m}_j = \mathbf{0}$. When the parameters are constrained to some known values $\mathbf{m}_j$, $\mathbf{C}_j$ is a unit matrix. Each constraint defines a pdf of model parameters (Jackson and Matsu'ura, 1985):

$$p\left(\mathbf{m}\middle|\sigma_j^2\right) = \frac{1}{\left(2\pi\sigma_j^2\right)^{N_j/2}}\left|\mathbf{C}_j^\mathrm{t}\mathbf{C}_j\right|_\mathrm{D}^{1/2} \exp\left(-\frac{1}{2\sigma_j^2}\left(\mathbf{m}_j - \mathbf{C}_j\mathbf{m}\right)^\mathrm{t}\left(\mathbf{m}_j - \mathbf{C}_j\mathbf{m}\right)\right) \quad [21]$$

where N_j is the number of independent constraint equations, namely, the rank of $\mathbf{C}_j$, and $|\bullet|_\mathrm{D}$ represents the product of all nonzero eigenvalues. When there are N_c constraint matrices, the joint pdf is given by a product of eqn [21] for $j = 1, \ldots, N_\mathrm{c}$ as

$$p\left(\mathbf{m}\middle|\sigma_1^2, \cdots, \sigma_{N_\mathrm{c}}^2\right) = A\exp\left(-\sum_{j=1}^{N_\mathrm{c}}\frac{1}{2\sigma_j^2}\left(\mathbf{m}_j - \mathbf{C}_j\mathbf{m}\right)^\mathrm{t}\left(\mathbf{m}_j - \mathbf{C}_j\mathbf{m}\right)\right) \quad [22]$$

where A is a normalization factor that keeps the integration of eqn [22] for $\mathbf{m}$ to unity.

Hereafter, we consider only one constraint matrix $\mathbf{C}_1$, whose rank is equal to N_m, for simplicity. Following the Bayesian theorem, a posterior pdf is obtained by multiplying eqn [21] and the likelihood function eqn [13]:

$$p(\mathbf{m},\sigma^2,\sigma_1^2|\mathbf{d}^0) \propto L(\mathbf{m},\sigma^2|\mathbf{d}^0)p(\mathbf{m}|\sigma_1^2)$$
$$\propto (\sigma^2)^{-N_d/2}(\sigma_1^2)^{-N_m/2}\exp\left(-\frac{\mathrm{Res}(\mathbf{m})}{2\sigma^2}\right) \quad [23]$$

$$\mathrm{Res}(\mathbf{m}) = (\mathbf{d}^0-\mathbf{G}\mathbf{m})^t\mathbf{E}_d^{-1}(\mathbf{d}^0-\mathbf{G}\mathbf{m}) + \alpha_1^2(\mathbf{m}_1-\mathbf{C}_1\mathbf{m})^t(\mathbf{m}_1-\mathbf{C}_1\mathbf{m}) \quad [24]$$

where, $\alpha_1^2=\sigma^2/\sigma_1^2$. For given σ^2 and σ_1^2, the problem is formally identical to the maximum likelihood problem in eqns [13]–[15], and the solution is similarly obtained by maximizing the posterior pdf eqn [23], in other words minimizing the least square residual $\mathrm{Res}(\mathbf{m})$. The best estimates and the covariance matrix of parameters are

$$\mathbf{m}_0 = \left(\mathbf{G}^t\mathbf{E}_d^{-1}\mathbf{G}+\alpha_1^2\mathbf{C}_1{}^t\mathbf{C}_1\right)^{-1}\left(\mathbf{G}^t\mathbf{E}_d^{-1}\mathbf{d}^0+\alpha_1^2\mathbf{C}_1{}^t\mathbf{m}_1\right) \quad [25]$$

$$\mathbf{E}_m = \left(\mathbf{G}^t\mathbf{E}_d^{-1}\mathbf{G}+\alpha_1^2\mathbf{C}_1{}^t\mathbf{C}_1\right)^{-1}\sigma^2 \quad [26]$$

To find the best estimates of σ^2 and σ_1^2, we integrate eqn [23] for the model parameters and obtain the marginal likelihood as

$$L(\sigma^2,\sigma_1^2) = \int p(\mathbf{m},\sigma^2,\alpha_1^2|\mathbf{d}^0)\mathrm{d}\mathbf{m} \propto (\sigma^2)^{-N_d/2}(\alpha_1^2)^{-N_m/2}\left|\mathbf{G}^t\mathbf{E}_d^{-1}\mathbf{G}+\alpha_1^2\mathbf{C}_1{}^t\mathbf{C}_1\right|^{-1/2} \times \exp\left(-\frac{\mathrm{Res}(\mathbf{m}_0)}{2\sigma^2}\right) \quad [27]$$

where $|\bullet|$ represents the determinant. At the maximum of the marginal likelihood, the variance of data error σ^2 is calculated as (Yabuki and Matsu'ura, 1992)

$$\sigma^2 = \frac{\mathrm{Res}(\mathbf{m}_0)}{N_d} \quad [28]$$

Following the definition of well-known Akaike's information criterion (AIC), Akaike (1980) introduced ABIC as

$$\mathrm{ABIC}(\alpha_1^2) = -2(\log \text{marginal likelihood}) + 2(\text{number of hyperparameters}) \quad [29]$$

For the current problem, the expression of ABIC is given by eqn [27] and eqn [28] as

$$\mathrm{ABIC}(\alpha_1^2) = N_d\ln(\mathrm{Res}(\mathbf{m}_0)) - N_m\ln(\alpha_1^2) + \ln\left|\mathbf{G}^t\mathbf{E}_d^{-1}\mathbf{G}+\alpha_1^2\mathbf{C}_1{}^t\mathbf{C}_1\right| + \mathrm{const} \quad [30]$$

The ABIC is a function of the hyperparameter α_1^2, and its minimum is found numerically.

When the constraint is too strong, α_1^2 is very large, and the second and the third terms of eqn [30] almost cancel, but the least-squares residual $\mathrm{Res}(\mathbf{m}_0)$ in the first term is also large and so is ABIC. When the constraints are too weak, α_1^2 is very small, the residual is small, but the second term becomes large to make a large ABIC. Between these two limits, we can find a minimum of ABIC. In eqn [30], $1/\alpha_1^2$ defines the condition number of the matrix in the parenthesis of the third term. The trade-off between the residual and the condition number of the matrix is essential for ABIC. Cohee and Beroza (1994) compared the residual and the condition number to find an appropriate value of weight of constraint. Several methods to estimate the appropriate weight of constraints are introduced in Aster et al. (2012). Equation [30] gives a statistical background for such a comparison. In other words, the ABIC solves the well-known problem of a trade-off between the resolution and accuracy of the solution. However, practically, finding the minimum of ABIC is not easy for general inversion problems including nonlinearity and non-Gaussian error distribution. Fukuda and Johnson (2008) demonstrated that the correct estimation of ABIC with frequently used nonnegative constraints is only possible through time-consuming Monte Carlo sampling.

We can use a combination of different constraints using eqn [22] instead of eqn [21]. The ABIC with two constraints is explained in Fukahata et al. (2004). **Figure 4** is a typical example of application of ABIC for slip inversion, with prior information that slip distribution is smooth in space and time. The synthetic waveforms from an input slip model (**Figure 4(a)**) contaminated with artificial noises are inverted for slip models (**Figures 4(b)–4(f)**) using different values of two hyperparameters in space and time, α_{space}^2 and α_{time}^2, respectively. The hyperparameters (weights of constraints) change the roughness of the slip distribution and source time functions. If the input model is consistent with the prior information and smooth to some extent, the slip model based on the minimum ABIC (**Figure 4(b)**) is similar to the input model.

4.09.3.3 Comparison in the Frequency Domain

The final goal of slip inversion is to obtain a model that can explain seismic waves over a broad-frequency band. Accurate strong-motion prediction requires such a broadband seismic source. However, most previous slip models are determined by mainly low-frequency (long-period) seismic waves. Not many studies discussed the validity of their model in the frequency domain. There have been a few studies that solved inversion problems in the frequency domain, such as Olson and Anderson (1988) and Cotton and Campillo (1995). The study of Ji et al. (2002a) using wavelet transform instead of Fourier transform is a mixed one in the time and frequency domains (wavelet domain). In this subsection, we discuss the nature of seismic waves at different frequencies and review studies that focused on the frequency contents of seismic waves.

A good reference model for seismic wave spectra is the omega square model (Aki, 1967; Brune, 1970), in which the displacement spectrum is proportional to

$$\frac{1}{1+\left(\frac{\omega}{\omega_0}\right)^2} \quad [31]$$

where ω_0 is a corner angular frequency. High-frequency waves attenuate as squared frequency, and lower-frequency waves relative to the corner frequency dominate in displacement records. Since waveform inversion in the time domain gives a solution that explains waves of dominant frequencies, the inversion of displacement data tends to explain these low-frequency waves. Since velocity data have large amplitude around the corner frequency, the model is determined by waves around this frequency. Similarly, acceleration data have implicit weights above the corner frequency. Naturally, records of seismometers with different instrumental responses (e.g., Mendoza and Hartzell, 1988a), or records with different band-pass filters (e.g., Ide, 1999), result in the apparently

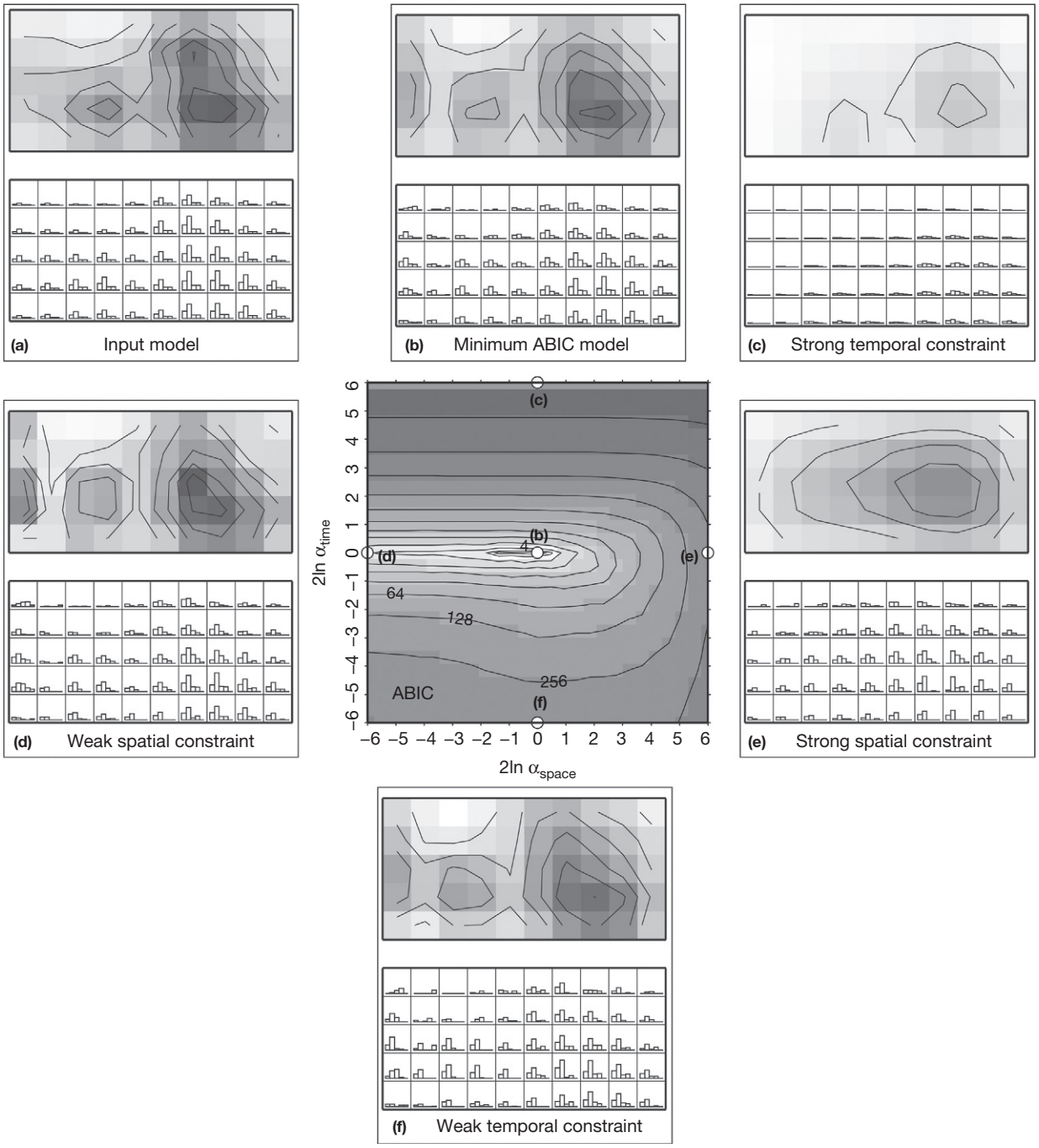


Figure 4 Simulated example of application of Akaike's Bayesian information criterion (ABIC) to slip inversion. Synthetic displacement at several stations surrounding the input source model are calculated and contaminated with noise of fractional Gaussian type. Slip models are determined for the different set of spatial and temporal hyperparameters, α^2_{space} and α^2_{time}, and the corresponding values of ABIC are shown in the central panel as a function of α^2_{space} and α^2_{time}. Since the absolute values are meaningless in this simulation, we set the origin of hyperparameters at the minimum of ABIC. (a) Input model: the slip distribution (upper panels) and time functions (lower panels). There are 10×5 subfaults and each subfault has five time windows consisting 250 parameters. (b) The model that provides the minimum ABIC. (c)–(f) Different models determined by stronger or weaker constraints.

different slip distributions. This is one reason of difference between models for the same earthquake, because different studies use different types (displacement, velocity, or acceleration) of data as shown in Table 1. To obtain a slip model that is valid over a broad-frequency band, we must impose some weighting schemes that weaken this omega square frequency dependence of seismic waves, which is difficult in the time domain.

Olson and Anderson (1988) proposed the first inversion in the frequency domain. They transformed time domain waves to frequency domain spectra and also converted Green's functions and slip distribution to the frequency–wave number domain. They compared data and predictions at each frequency, which greatly simplifies the problem especially at low frequencies and small wave numbers, preserving its linearity. For high frequencies and large wave numbers, the number of parameters increases and some regularization is required. They proposed the minimum length regularization, but did not apply it to the real data. Mendez and Anderson (1991) applied essentially the same method to the 1985 Michoacan earthquake (Mw 7.9) with a different regularization that minimizes an error function based on the discretized version of a wave equation, which is minimized by a propagating rupture at a constant velocity. One shortcoming of the frequency–wave number domain transform of slip distribution is that the positivity constraint is difficult to apply because the slip at a point is represented by the summation of slip functions at various frequencies and wavelengths.

Cotton and Campillo (1995) compared Fourier coefficients of data and model prediction in the frequency domain, but they modeled slip distribution in the time domain, in similar form to the nonlinear expression in **Figure 2(b)**. They applied this method to the real data of the Landers earthquake and obtained a slip model that is consistent with other studies. Actually, their method is essentially the same as an ordinary time domain inversion that compares the time domain data $\mathbf{d}^0$ and corresponding model predictions $\mathbf{d}^e$ and minimizes the least-squares residual $(\mathbf{d}^0-\mathbf{d}^e)^t(\mathbf{d}^0-\mathbf{d}^e)$. This value is an invariant for any orthonormal transformation T:

$$\left(\mathbf{T}\mathbf{d}^0-\mathbf{T}\mathbf{d}^e\right)^t\left(\mathbf{T}\mathbf{d}^0-\mathbf{T}\mathbf{d}^e\right)=\left(\mathbf{d}^0-\mathbf{d}^e\right)^t\left(\mathbf{d}^0-\mathbf{d}^e\right) \qquad [32]$$

Since the Fourier transform is a kind of orthonormal transformation, the comparison of Fourier coefficients in the frequency domain gives an identical residual to the one in the time domain. This property is also known as Parseval's theorem. Nevertheless, it also suggests the possibility to control relative weights of observed data in the frequency domain.

Although the Fourier transform is popular and its properties are well known, it has an aspect that is not easy to deal with. Fourier transform of a boxcar function in the time domain extends over infinite frequency band, and vice versa. In contrast, orthonormal wavelet transform is a transform between a finite time domain signal and a finite frequency domain spectrum. Ji et al. (2002a) adopted Meyer–Yamada wavelet transform to account for the contribution of finite frequency signal in slip inversion. Their method enables easy reweighting in each frequency band in the wavelet domain. They applied this method to the Hector Mine earthquake (Mw 7.1; Ji et al., 2002b).

While constructive wave interference occurs at low frequencies, high-frequency seismic waves are incoherent and the wave energies, squared amplitudes, are summed up at observation points. Special methods have been proposed to invert high-frequency incoherent waves and find the origin of these waves. A class of slip inversion developed for high-frequency waves is the envelope inversion. Synthetic envelopes from a source model are calculated using isochrone approximation of Spudich and Frazer (1984) (Zeng et al., 1993), using eGf (Hartzell et al., 1996; Kakehi and Irikura, 1996, 1997; Kakehi et al., 1996), and based on the scattering theory of Sato et al. (1997) and Nakahara et al. (1998, 1999). In many cases, the sources of high-frequency waves seem to be located near the periphery of large slip, which suggests high-frequency wave generation due to rapid accelerations and decelerations of the rupture front (Madariaga, 1977). However, many exceptions exist for this observation and there is no general relation between low-frequency slip distribution and high-frequency energy radiation. Apparently, this is a field that requires much further investigation.

4.09.3.4 Back-Projection Method: An Alternative to Slip Inversion

Slip inversion visualizes earthquake rupture process in a broadband frequency range. Focusing in a narrow frequency range, we can use another kind of visualizing method utilizing the coherence of seismic waves simultaneously observed in numerous seismometers, which has become available recently owing to the development of dense seismometer networks such as Hi-Net, Japan (http://www.hinet.bosai.go.jp), and the USArray seismometer network (http://www.earthscope.org/observatories/usarray). This is the back-projection method, firstly introduced by Ishii et al. (2005) to resolve the rupture process of the 2004 Sumatra earthquake. Unlike slip inversion that requires often subjective assumptions on Green's functions and constraints, the assumption for the back-projection method is relatively simple. Here, we review a typical treatment for this method.

We consider a seismic source that consists of K point sources, each of which has a common focal mechanism located at $\boldsymbol{\xi}^k$ and a slip time function $\Delta u(\boldsymbol{\xi}^k, t), k=1,\ldots,K$. The teleseismic P waves in the vertical displacement observed at location $\mathbf{x}^l, l=1,\ldots,L$ are approximated using the summation of contribution from all sources (Aki and Richards, 2002) as

$$u(\mathbf{x}^l,t)\sim C\sum_{k=1}^{K}\Delta\dot{u}\left(\boldsymbol{\xi}^k,t-T\left(\mathbf{x}^l,\boldsymbol{\xi}^k\right)\right) \qquad [33]$$

where C is a constant representing the factor controlled the radiation pattern and distance from the source. $T(\mathbf{x}^l,\boldsymbol{\xi}^k)$ is the travel time of P wave from $\boldsymbol{\xi}^k$ to $\mathbf{x}^l$. The success of the back-projection analysis depends strongly on the accuracy of this travel time. Technically, it is often given as the theoretical travel time adjusted using a station correction, which can be estimated accurately using the correlation between the initial part of waveforms (Ishii et al., 2005).

Focusing at a source location $\boldsymbol{\xi}^X$ and shifting time by $T(\mathbf{x}^l, \boldsymbol{\xi}^X)$, eqn [33] is modified as

$$\begin{aligned}u\left(\mathbf{x}^l,t+T\left(\mathbf{x}^l,\boldsymbol{\xi}^X\right)\right)\sim{}&C\Delta\dot{u}\left(\boldsymbol{\xi}^X,t\right)\\&+C\textstyle\sum_{k\neq X}\Delta\dot{u}\left(\boldsymbol{\xi}^k,t-\Delta T\left(\mathbf{x}^l,\boldsymbol{\xi}^k,\boldsymbol{\xi}^X\right)\right)\end{aligned} \qquad [34]$$

where $\Delta T(\mathbf{x}^l, \boldsymbol{\xi}^k, \boldsymbol{\xi}^X)=T(\mathbf{x}^l, \boldsymbol{\xi}^k)-T(\mathbf{x}^l, \boldsymbol{\xi}^X)$. The previous equation is written for each of L array stations. While the first term of eqn [34] is common for all stations, the second term is different and approximated to be random. Therefore, stacking eqn [34] for all stations will enhance the first term relative to

the second term. When L is large and the second term is negligible, the slip rate time function of a point source located at $\boldsymbol{\xi}^X$ is approximated as

$$\Delta \dot{u}\left(\boldsymbol{\xi}^X, t\right) \sim \frac{1}{CL}\sum_{l=1}^{L} u\left(\mathbf{x}^l, t + T\left(\mathbf{x}^l, \boldsymbol{\xi}^X\right)\right) \qquad [35]$$

This is the basic equation for the back-projection method. Simply computing eqn [35] for all possible locations of the source, we obtain back-projection images of earthquake rupture, which is proportional to the slip rate.

To neglect the second term, $\Delta \dot{u}(\boldsymbol{\xi}^X, t)$ must have a short duration enough that the waves from different source locations do not correlate. Otherwise, we observe some interference from the second term. Although the real slip function has some finite rise time, it is approximated as a sequence of pulse-like slip, in a narrow frequency range. Therefore, the approximation of eqn [35] is valid only in a narrow high-frequency range, where the finiteness of the rise times is not visible. Usually, back-projection image is obtained from high-pass-filtered waveforms or waveforms recorded by a short-period seismometer array like Hi-Net (Ishii et al., 2005).

Practically, eqn [35] is replaced by an efficient stacking algorithm like Nth root stack (Rost and Thomas, 2002), and some weighting schemes among stations are also applied based on the station distribution and S/N ratio. Instead of stacking waveforms, we may determine the direction of incoming waves into a seismometer array at specific timing, by methods such as the MUSIC method (Meng et al., 2011) and the compressive sensing (Yao et al., 2011). The back-projection method may be applied for phases other than P waves. Kiser et al. (2011) demonstrated that the back-projection method applied to a combination of the direct and depth phases improves the resolution and enables discussion of the depth extent of rupture at intermediate depth. For shallow depth earthquakes, depth phases are difficult to separate and can violate the approximation of eqn [33]. Yagi et al. (2012) showed that convolution with theoretical Green's function reduces the artifact due to depth phases.

4.09.4 Example of Slip Model

4.09.4.1 The 1999 Chi-Chi Earthquake

The Chi-Chi, Taiwan, earthquake is an interplate thrust earthquake between the Eurasian and the Philippine Sea plates, which occurred on 21 September 1999 (local time). Surface offsets appeared along an 85 km segment of the Chelungpu Fault. The maximum offset of about 8 m is located near the northern end where the fault trace becomes complex and changes the strike direction to the east (**Figure 5**). Aftershocks are distributed to the east of this surface trace, reflecting the east dipping fault plane. Digital strong-motion records are available from 441 stations, 60 of which are within 20 km from the surface trace. These stations are maintained by the Taiwan Strong-Motion Instrumentation Program (Lee et al., 2001). Furthermore, coseismic static displacements at about 100 sites are available from GPS surveys (Yu et al., 2001). They comprised the best data set in quality and quantity among all the earthquakes before 2005.

Several groups determined a spatiotemporal slip model using this data set. Chi et al. (2001) assumed one fault plane and applied a multi-time-window method to invert strong-motion velocities and surface offsets. Wu et al. (2001) used spatiotemporal basis functions and Bayesian modeling to jointly invert strong-motion velocities and GPS displacements. They assumed from one to three fault planes to explain the northern fault bend. Ma et al. (2001) used a multi-time-window method for strong-motion displacements, far-field displacements, and GPS data. They compared two analyses with and without a northern small fault striking to the northeast. The northern small subfaults assumed by Wu et al. (2001) and Ma et al. (2001) were necessary to explain the records at some stations close to the surface offsets, but they are almost independent of the rest of the data. All three studies used linear parameterization and solved using NNLS. The large number of parameters from 4000 to 8000 requires regularization to get reasonable solutions. On the other hand, Zeng and Chen (2001) inverted strong-motion and GPS data to determine the distribution of slip, rupture time, and rise time in a nonlinear expression using a GA. Their fault plane has a 3-D geometry, an extension of the surface trace to the eastward down. Ji et al. (2003) applied their wavelet transform inversion method with SA to resolve slip distribution on a 3-D complex fault plane.

Figure 5 shows the locations of the assumed fault planes, final slip distributions, and rupture propagation behaviors in Chi et al. (2001), Ma et al. (2001), and Wu et al. (2001) (models C, M, and W, respectively). Models M and W are the single fault models in these papers, and the slip data of models C and M are taken from the supplemental CD of the original papers. Although the values of seismic moment are written as $(2.7–4.7) \times 10^{20}$ Nm in the original paper, all the seismic potencies calculated from the model parameters are within $(1.0–1.15) \times 10^{10}$ m^3, which is reasonable because the colors of final slip in **Figure 5** are apparently similar.

The common features visible in these models are (1) total potency and rupture area, (2) small slip around the hypocenter, (3) large and deep slip in the northern part, (4) small and shallow slip in the southern part, and (5) the maximum slip of about 20 m. The good agreement between models M and W is probably due to addition of GPS data. In model C, the slip area extending over the northern end of the surface trace is hard to explain. The previous features from (1) to (5) are also found in the models of Zeng and Chen (2001) and Ji et al. (2003).

The most prominent similarity among the slip histories of these models is the large slip at 40 km north of the hypocenter and 20–24 s after the hypocentral time. The rupture propagation velocity is less than 2 km s^{-1}, which is slower than the average shear wave velocity in the source area of 2.5–3.0 km s^{-1}. This is also shown by Zeng and Chen (2001) and Ji et al. (2003). However, there are considerable discrepancies such as the timing of slip propagating to the south. Slip amount within each period does not match well. In **Figure 5**, the origin time of model M is shifted 2 s to match other models better. This difference is not small, but it is not unrealistic because the strong-motion records used in this study do not have accurate timing information and researcher independently corrected them based on other information.

When large amounts of data are available as the Chi-Chi earthquake, we observe similarities among models, which are

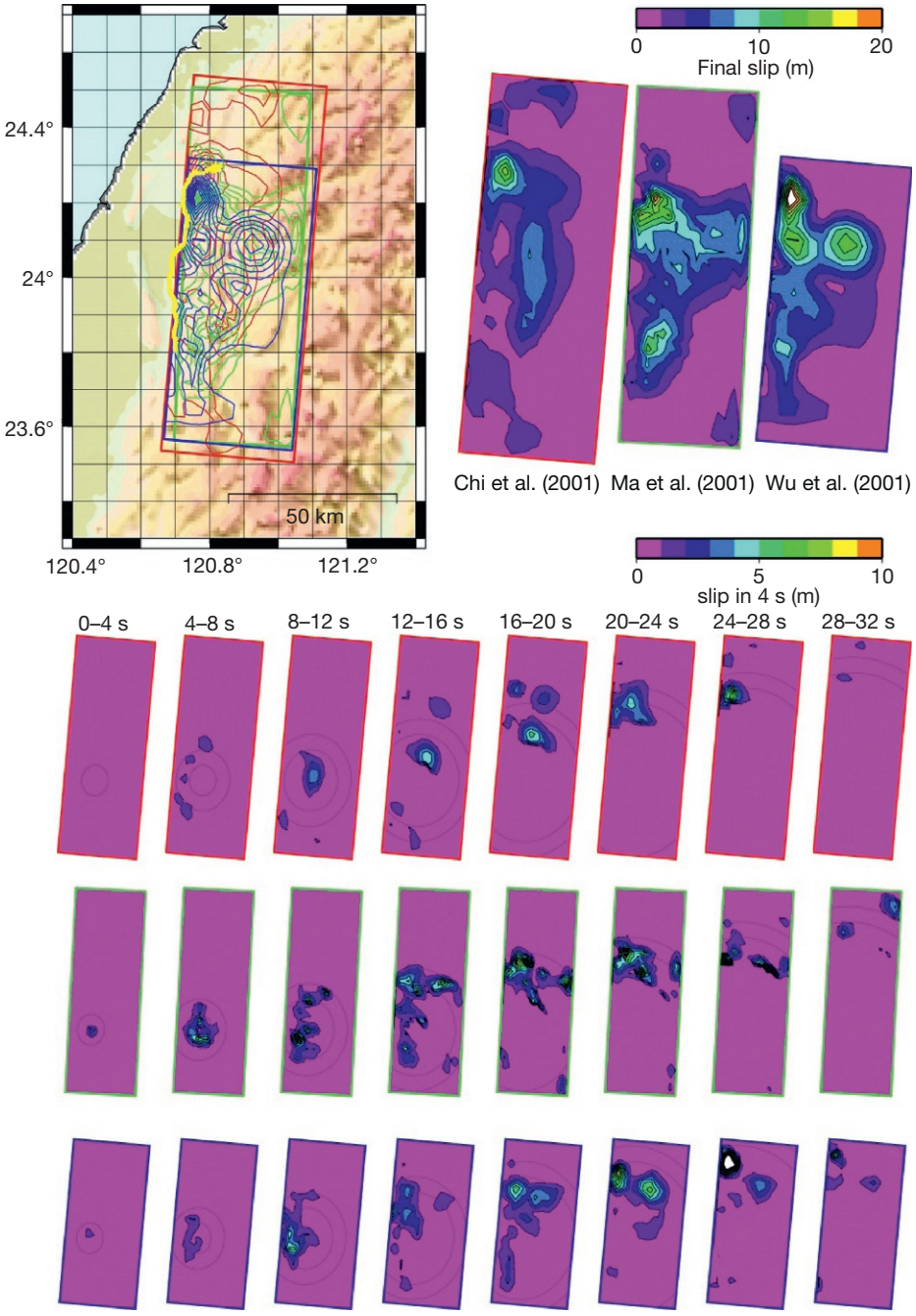


Figure 5 Comparison of three slip models by Chi et al. (2001), Ma et al. (2001), and Wu et al. (2001). All distribution is shown as the projection to the free surface. (Top left) Location in the map. Surface offsets along the Chelungpu Fault are shown by a yellow line. (Top right) Final slip distributions. (Bottom) Slip distributions during each time period.

certainly constrained by data regardless of slightly different model assumptions. A dense GPS network is especially useful to constrain the spatial distribution of slip. The amount and timing of shallow slip cannot be very different when the measurements of surface offsets and very near-fault strong-motion records are available. While very near-fault records are only sensitive to the local slip, a good coverage of seismic data is essential to reveal spatiotemporal rupture propagation. All these data are available for the Chi-Chi earthquake. Similarly, we can find common timing and spatial extent of large slip in the models of well-studied events with sufficient data, such as the Landers, Northridge, Kobe, and Tottori earthquakes (Table 1). Nevertheless, we should also acknowledge that considerable differences exist in the details of slip model that is affected by slight difference in the way of analysis even with such good data set. The behavior of small slips, often including those in the initial stage of rupture, is difficult to resolve.

4.09.4.2 The 2011 Tohoku-Oki Earthquake

The Tohoku-Oki earthquake (Mw 9.0) on 11 March 2011 was another well-recorded interplate earthquake. Although most observation points are located landward, which is a worse condition than the case of the Chi-Chi earthquake, it occurred close to Japanese dense observation networks. From the aspect of the variety and volume of data, it is so far the best-recorded earthquakes in the history of seismology. There have been numerous research papers on the source process of this event, and attempting to cover all of them exceeds the scope of this chapter. Here, I refer to the relatively early works. Data used for slip inversion include records of far-field broadband seismometers (e.g., Ide et al., 2011; Shao et al., 2011), near-field strong-motion seismometers (e.g., Suzuki et al., 2012), and onshore GPS (e.g., Iinuma et al., 2012; Ozawa et al., 2011). GPS is used as static displacements or an alternative of seismometers (e.g., Yue and Lay, 2011). Offshore GPS, seafloor pressure gauges, and tide gauges are also used to model tsunamis (e.g., Fujii et al., 2011). Many studies applied joint inversion for fault slip, using several types of data (Koketsu et al., 2011; Romano et al., 2012; Simons et al., 2011; Yoshida et al., 2011). In addition, there are several studies applying back-projection methods, by the original version (Ishii, 2011; Koper et al., 2011; Wang and Mori, 2011) and by some modified versions (Meng et al., 2011; Yagi et al., 2012; Yao et al., 2012). The attempt to determine the 'most reasonable' model of the Tohoku-Oki earthquake is still ongoing.

Among many studies, **Figure 6** compares three slip models and one back-projection image model. The slip models are the results by an eGf inversion using teleseismic P waves (Ide et al., 2011), theoretical Green's function analysis using near-field strong-motion data alone (Suzuki et al, 2012), and joint slip inversion with near-field strong-motion data, GPS, and tsunami waveforms (Yokota, 2013). The slip distributions in six 20 s time windows, up to 120 s, are shown with a unified scale and colors. The back-projection images from the modified back-projection method determined using the teleseismic waves between 0.5 and 2.0 Hz (Yagi et al., 2012) are also compared.

There are common features among previously published slip models, including models in **Figure 3**. The first rupture up to 40 s radiated a distinct pulse from a source region around the hypocenter and deeper, with the rupture front propagating downward traced by the back-projection images. The second rupture was the largest, near the trench at about 60–80 s. The estimated maximum slip at the trench ranges from 30 to 80 m, though the maximum is highly dependent on the smoothing conditions. While slip inversions in broadband frequency range clearly resolve this large slip, it is not visible at all in the high-frequency back-projection images. This rupture was

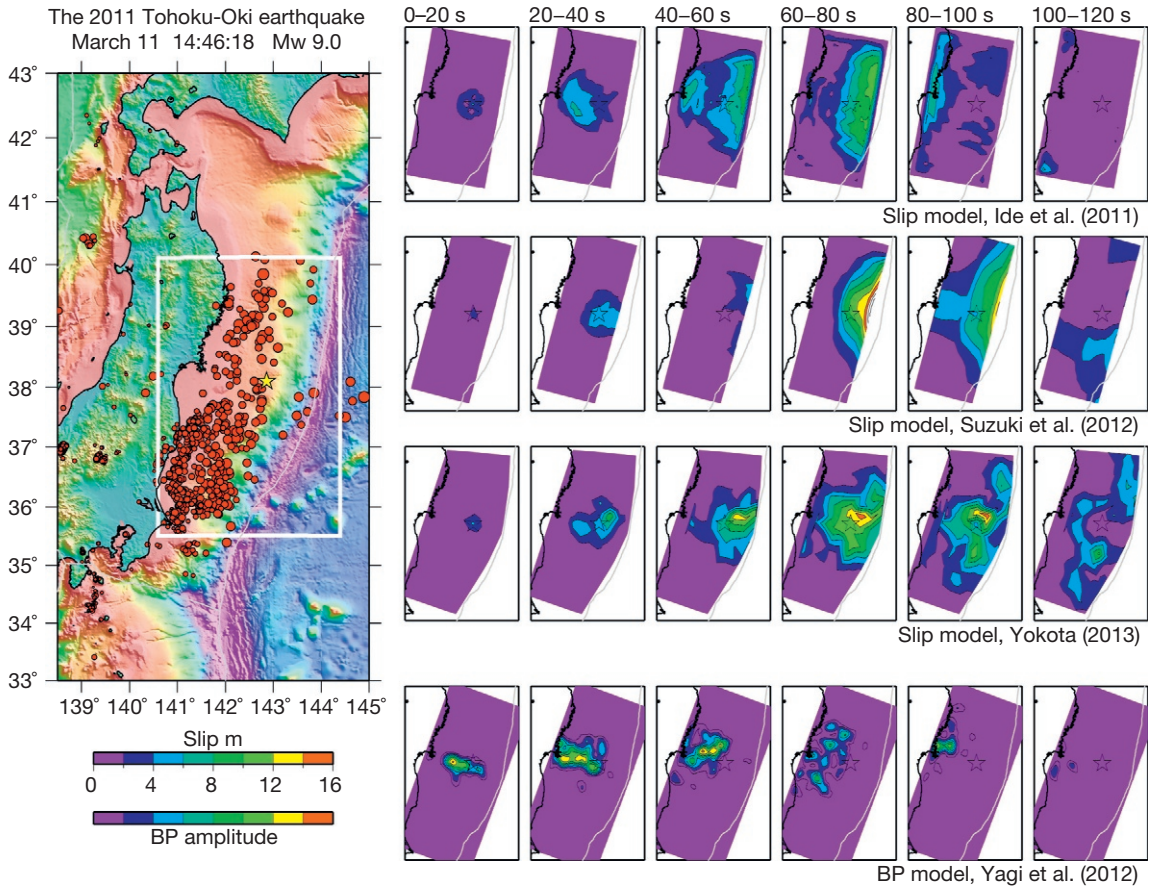


Figure 6 Comparison of three slip models by Ide et al. (2011), Suzuki et al. (2012), and Yokota (2013) and a back-projection model by Yagi et al. (2012). All distribution is shown as the projection to the free surface. (Left) Location in the map. A yellow star and red circles are the mainshock hypocenter and aftershock by March 12, respectively. White box shows the area to show the rupture history.

followed by another rupture at deep region, which likely overlapped with the first rupture region, namely, some part of the plate interface likely to have ruptured at least twice. Most of the seismic moment was released in 100 s, and afterward, some small ruptures occur in the southern part of the plate interface.

Despite the similarities of these characteristic features, the timing of the large slip is somewhat biased between models in Figure 3. Potential sources of these discrepancies are the inaccuracy of Green's functions and misalignment between observed and theoretical data. The former must be a problem in empirical analysis of Ide et al. (2011), because the variation of depth phases is not well modeled by a single event, although the timing of P wave arrivals is accurately aligned between the main and eGf events. On the other hand, the time alignment between observed and theoretical waveforms is a difficult task for strong-motion inversions, especially for the Tohoku-Oki earthquake, because the amplitude of the initial P wave is very small and the azimuthal coverage of stations is basically one-sided (west). Yokota (2013) used 3-D Green's functions, but the characteristics of their model are fairly similar to their 1-D model (Koketsu et al., 2011). Both of these models are different from models of Suzuki et al. (2012), though all models depend on similar set of near-field strong-motion records to constrain the temporal evolution. This comparison suggests that the time alignment not explicitly presented in papers is very influential on the slip models.

The back-projection images also have been determined by many groups, and the diversity of models relative to the reference point (assumed hypocenter location) seems to be much smaller than that of slip models, suggesting general robustness of the methods. These methods also require the alignment of timing between observed waveforms, but it is performed objectively using cross correlations and not a serious problem. The shortcomings are that the information is limited in relatively high frequencies and that the estimated values do not have an explicit relation to physical quantities. Nevertheless, the combination of slip models and back-projection models provides vivid images of rupture process and enables us to speculate about the physical process behind the models.

4.09.5 Extended Studies Based on Slip Models

4.09.5.1 Characteristics of Slip Models

There are qualitative features frequently pointed out in many slip models, for example, pulse-like slip propagation, few aftershocks in the area with large slip, and the tendency of rupture initiation from the bottom of the fault plane. In this section, we review the discussion and physical interpretation of these features.

In many slip models, slip propagates like a pulse with a narrow width compared to the size of whole fault plane. Summarizing seven slip models, Heaton (1990) concluded that the slip pulse is a general feature of seismic rupture process and not an artifact due to the setting of time windows or limited frequency contents of inverted data. One good example among later studies is the Landers earthquake that ruptured 65 km faults, but the width of the slip area at any moment is less than 15 km (e.g., Wald and Heaton, 1994). Heaton (1990) proposed that the velocity-weakening nature of rock friction is responsible for pulse-like behavior. Cochard and Madariaga (1994) confirmed the generation of a slip pulse using velocity-weakening friction law in numerical simulation of dynamic crack propagation. However, this is not the only mechanism of slip pulse. Alternatively, Beroza and Mikumo (1996) and Peyrat et al. (2001) demonstrated that spatial heterogeneity of stress can produce slip pulses without a velocity-weakening friction law.

A reliable gross feature of slip inversion is the location and time of the largest slip. The location and time of rupture initiation (hypocenter) are far more reliable than that of the largest slip. Therefore, comparing these numbers, we can estimate overall rupture velocity. In the 1970s, rupture velocity was discussed using macroscopic earthquake models. Geller (1976) compiled rupture velocity of many macroscopic models and gave a typical rupture velocity as 72% of the S wave velocity, with a caution that the value depends on the model assumption. The result of slip inversion is useful to measure model-independent rupture velocity. The rupture velocity of 2.2 km s^{-1} in the Chi-Chi earthquake (Section 4.09.4) is a well-determined example. It is easy to measure a rupture velocity when there is a slip pulse in the model. In the model of Wald and Heaton (1994), the Landers earthquake propagated at 2.5–3.0 km s^{-1}, with deceleration near fault steps. The rupture velocity of the Sumatra earthquake is also fixed well as 2.8 km s^{-1} from the spatiotemporal energy radiation determined by Ishii et al. (2005). In all models, average rupture velocity is less than S wave velocity but a little faster than 70% suggested from macroscopic source models.

Unlike the average rupture velocity, local rupture velocity can exceed S wave velocity, although it is not robust because of smoothing introduced in slip inversion. In the slip model of the Imperial Valley earthquake based on trial and error, Archuleta (1984) showed that local rupture velocity is faster than S wave velocity. Frankel (2004) reported that the local rupture velocity is as fast as 5 km s^{-1}, while the average is about 3 km s^{-1} in the case of the 2002 Denali fault, Alaska, earthquake (Mw 7.9). Such local super shear rupture is expected when stress is heterogeneous (Day, 1982a) or fracture energy is heterogeneous with homogeneous stress (Ide and Aochi, 2005). However, the observational evidence is not strong enough to conclude that such super shear rupture propagation is a general phenomenon of earthquakes.

Estimation of rupture velocity for small earthquakes is difficult. Ide (2001) determined slip models of 18 moderate (M4–5) earthquakes that occurred in the Hida Mountains, Japan, but the estimation of rupture velocity has large uncertainties because of poor resolution. In the smallest slip model ever determined for earthquakes, the model for the micro (M0.8–1.4) earthquake in South Africa gold mine, Yamada et al. (2005) found that the rupture velocity is not less than 65% of S wave velocity. Although it is not a result based on slip models, Imanishi et al. (2004) found high rupture velocity for small (M1.3–2.7) earthquakes in western Nagano, Japan, using the method utilizing stopping phases (Imanishi and Takeo, 2002). These studies suggest that the rupture velocity is similar for earthquakes from M1 to M9 but each estimation was subject to large errors and obviously further studies are necessary.

Slip models are frequently compared with aftershock distributions. It is often claimed that aftershocks are rare near or in the areas on the fault where the slip was large. Mendoza and Hartzell (1988b) investigated seven slip models of earthquakes

in the western United States that are determined using near-field records and hence relatively reliable. Although it was qualitative, they showed that aftershocks occur mostly outside of or near the edges of the source areas indicated by the patterns of mainshock slip. Das and Henry (2003) investigated various slip models for many events including those in subduction zones and ridges. They found that few, and usually the smaller, aftershocks occur in the high-slip regions of the fault, with only one exception. Wiemer and Katsumata (1999) discussed slip models and b value of the Gutenberg–Richter frequency–magnitude relation (Gutenberg and Richter, 1944). For the 1984 Morgan Hill (Mw 6.2), Landers, Northridge, and Kobe earthquakes, they found that large slip regions seem to correspond to large b values. However, it should be noted that the large slip region has few aftershocks and the estimations of b value in most of these regions were impossible. Although the mechanism of aftershock genesis is not fully understood, the existence of many aftershocks surrounding the areas of large slip indicates the importance of the stress redistribution by the mainshock.

Another often reported behavior is rupture initiation from the bottom of the fault (e.g., Das and Scholz, 1983). One may easily compare the aftershock distribution and the location of mainshock hypocenter, but the comparison between slip distribution and the initial rupture location is more informative. From the large collection of published slip models, Mai et al. (2005) compared the location of initial rupture and slip distribution and concluded that the rupture started from near the bottom of the assumed fault plane in strike-slip and crustal dip-slip earthquakes. Although Das and Henry (2003) reported that the rupture tends to nucleate in the region of low slip or at the edge of high-slip regions, Mai et al. (2005) did not find such tendency.

Rupture direction of large earthquakes tends to be unidirectional as shown by McGuire et al. (2001) using a 2nd moment inversion method for far-field records. We can estimate the 2nd moment for slip models and determine overall rupture direction. McGuire et al. (2002) investigated slip models for earthquakes larger than M7 and confirmed his earlier conclusion based on far-field observation.

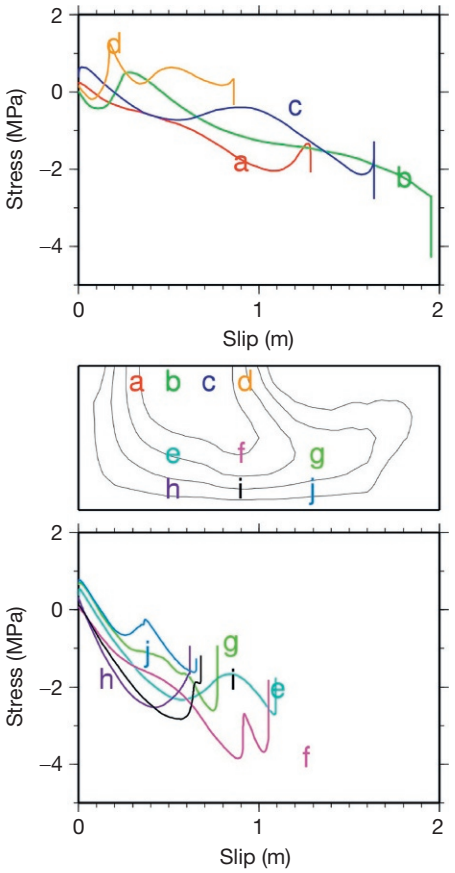


Figure 7 The relation between slip and stress on the fault plane, determined for the 1995 Kobe earthquake (Ide and Takeo, 1997). Each trace is the function calculated at the corresponding location in the middle figure that shows final slip distribution by contours. Reproduced from Ide S and Takeo M (1997) Determination of constitutive relations of fault slip based on seismic wave analysis. *Journal of Geophysical Research* 102: 27379–27391.

4.09.5.2 Implication of Slip Models for Fault Dynamics

Spatiotemporal slip distribution is directly connected with spatiotemporal stress distribution. Once spatiotemporal slip distribution is obtained, it is used as a boundary condition for calculation of stress change during the rupture propagation (Bouchon, 1997; Ide and Takeo, 1997). This yields a slip–stress relation at each point on the fault plane. **Figure 7** is the first result of such calculation for the Kobe earthquake (Ide and Takeo, 1997). The slip–stress relation is basically slip-weakening, similar to the theoretical assumption of Ida (1972) and Andrews (1976). The slip-weakening distance Dc is estimated as 50 cm to 1 m. However, it was also shown by Ide and Takeo (1997) that Dc less than 50 cm is not detectable for the Kobe earthquake because of the limited resolution. Olsen et al. (1997) also estimated Dc of 0.8 m for the Landers earthquake using dynamic modeling. However, this estimation might also suffer from the effect of the limited resolution. Guatteri and Spudich (2000) demonstrated that similar seismic waves are radiated from dynamic models of different Dc distribution, as far as they have similar distribution of fracture energy Gc, which scales with the product of Dc and breakdown stress drop (i.e., the difference between yield stress and dynamic friction level). The estimation of Gc is robust compared to that of Dc because Gc is controlled by rupture velocity and less affected by the limited resolution. Therefore, the most stable dynamic constant derived from slip inversion is not Dc, but Gc.

Beroza and Spudich (1988) first estimated Gc as 2 MJ m^{-2} from the slip model of the Morgan Hill earthquake. Guatteri et al. (2001) applied a rate- and state-dependent friction law for the model of Ide and Takeo (1997) and obtained a value of Gc as 1.5 MJ m^{-2}. In addition to the fracture energy Gc measured from slip–stress relation, radiated energy from earthquake Es is also calculated using the expression of Kostrov (1974) as

$$E_s = \frac{1}{2}\int_\Sigma \left(\sigma_{ij}(\mathbf{x},0) - \sigma_{ij}(\mathbf{x},\infty)\right)\Delta u_i(\mathbf{x},\infty)\nu_j \mathrm{d}S - \int_\Sigma \gamma(\mathbf{x})\mathrm{d}S + \int_0^\infty \mathrm{d}t \int_{\Sigma(t)} \dot{\sigma}_{ij}(\mathbf{x},t)\Delta u_i(\mathbf{x},t)\nu_j \mathrm{d}S \quad [36]$$

where $\sigma_{ij}(\mathbf{x}, t)$ is the stress on the fault plane and $\gamma(\mathbf{x})$ is the free surface energy to make an infinitesimally small slip on the fault

plane. Summation convention is used. The third term contains the fracture energy derived from the slip–stress relation, which is also called seismological fracture energy and breakdown work (Cocco et al., 2006). Ide (2002, 2003) applied this equation to slip models of several earthquakes to estimate the seismic energy and fracture energy. While the fracture energy is in the order of MJ m^{-2} for the Kobe earthquake, total radiated energy calculated by this method is smaller by a factor of 3 than the values determined using far-field body waves (Boatwright and Choy, 1986). This suggests that the slip model does not account for the total energy radiation because it is determined by seismic waves of relatively lower frequencies with smoothing constraints.

Another method to estimate the slip-weakening distance Dc was proposed by Mikumo et al. (2003) and Fukuyama et al. (2003). Instead of the usual definition of Dc in a slip–stress diagram, they measured the slip amount when slip rate reaches the maximum as Dc′. Applying this method to the 2000 Tottori and the 1995 Kobe earthquakes, they concluded that Dc′ scales with the final slip. However, Spudich and Guatteri (2004) suggested that it can be an artifact due to low-pass filters. In fact, when the frequency contents of the data have an upper limit, that limit affects the maximum value of Dc′. Therefore, as Spudich and Guatteri (2004) pointed out, it is important to compare data and model prediction in the frequency domain as reviewed in **Section 4.09.3.3**.

If fault plane is flat, absolute stress level is usually irrelevant to seismic wave radiation. However, when the rake of fault slip rotates during the rupture process and frictional traction is collinear with the instantaneous slip motion, there is a unique absolute stress consistent with the motion (Spudich, 1992). Some slip models allow rake angles to vary during slip, which can be used to estimate the absolute stress level (Guatteri and Spudich, 1998; Spudich, 1992). Since the details of slip distribution from slip inversion depend on researcher's assumptions, we should be careful about judging whether the observed rake rotation is constrained directly by the data and not an artifact due to model assumptions.

4.09.5.3 Dynamic Modeling and Slip Models

Most slip models are kinematic models and some of the features are not physically realistic. For example, stress diverges to infinity around subfault boundaries with a constant slip, and dynamically feasible slip function is usually asymmetrical unlike the frequently used triangle function (**Figure 2(a)**). In the previous subsection, we reviewed dynamic implications directly deduced from such kinematic slip models. On the other hand, dynamic modeling is performed to construct an elastodynamically reasonable source model using only some aspects of slip inversion, for example, final slip distribution or rupture timing. A dynamic source model is constructed based on fracture mechanics, with the assumption of stress state and friction/fracture law. This is one practical way to discuss source behavior and seismic wave radiation at high frequencies where kinematic slip models are no longer reliable.

The problem of crack propagation in elastic medium has been solved numerically since the late 1970s (Andrews, 1976; Das and Aki, 1977; Day, 1982b). While early studies solved some ideal cases, the first application to a complex rupture process of real earthquake was by Quin (1990) who carried out dynamic modeling of the Imperial earthquake based on the slip model of Archuleta (1984). He calculated the stress drop from the slip distribution of the slip model, assumed the distribution of strength, and simulated spontaneous crack propagation at a prescribed rupture velocity of the original model satisfying the fracture criterion of the maximum stress (strength). Miyatake (1992) applied a similar method to four inland earthquakes in Japan. Since these procedures use only final slip distribution and rupture timing of the original slip model, seismic waves radiated from the dynamic model cannot explain the observation. Fukuyama and Mikumo (1993) and Ide and Takeo (1996) developed dynamic modeling methods that iteratively improve models to explain observed waveforms. The earlier mentioned studies of dynamic modeling assume the maximum stress fracture criterion (Das and Aki, 1977), which depends on the grid size. Rather, natural fault slips must obey a friction law. Olsen et al. (1997) carried out the first dynamic modeling to simulate the Landers earthquake, in which slip occurs following a slip-weakening friction law. This dynamic model is later improved by Peyrat et al. (2001) to determine the heterogeneous stress distribution on a flat fault plane. Dynamic modeling using an elliptical patch has been developed by Ruiz and Madariaga (2011) and applied to the Mw 6.7 Michilla, Chile, earthquake.

One of the recent major topics in dynamic modeling is the geometry of fault planes. Unlike a strike-slip fault, a shallow dipping fault does not have symmetry because of the existence of the free surface. Oglesby et al. (1998) carried out dynamic modeling of a 2-D shallow thrust fault and showed that the change of normal stress on the fault has an important effect on the fracture criterion at the rupture front. This result suggests that a complex slip distribution that implies stress heterogeneity can be due to complex fault geometry. Actually, Oglesby and Day (2001) showed that homogeneous stress and dipping fault plane can explain the gross feature of observed near-source ground motion for the Chi-Chi earthquake, for example, the larger movement in the hanging wall than the footwall. Dynamic rupture propagation along a more complex fault system has become feasible. Aochi and Fukuyama (2002) showed that homogeneous background stress and change of strike angles in a fault system are essential to explain spontaneous rupture transfer between three major faults during the Landers earthquake. Aochi et al. (2003) also showed that the stress distribution of Peyrat et al. (2001) improves the details of dynamic slip distribution.

4.09.5.4 Scaling of Earthquake Heterogeneity

Most scaling relations between macroscopic parameters of earthquake source are derived from a geometric similarity, that is, constant stress drop scaling (e.g., Kanamori and Anderson, 1975). However, the breakdown of the scaling relations is not well understood. There are still debatable issues, such as the effect of finite width of seismic layer on the scaling between fault length and slip amount (e.g., Matsu'ura and Sato, 1997; Scholz, 2002) and size dependence of the ratio between seismic energy and seismic moment, especially for small earthquakes (e.g., Ide and Beroza, 2001; Kanamori and Brodsky, 2001).

On the other hand, studies have recently started for the scaling of the complex earthquake source. It is not straightforward to extract meaningful parameters from various slip models. Using many published slip models, Somerville et al. (1999) determined the size of fault plane excluding the regions of little slip near the edges. Furthermore, they defined an asperity as "a region on the fault rupture surface that has large slip relative to the average slip on the fault," measured the total area of rectangular asperities in each model, and concluded that the asperities occupy 17.5% of the fault plane. The fault size of the slip model is also measured by Mai and Beroza (2000) using an autocorrelation function. They obtained a nonlinear relation between slip and fault length for large strike-slip events. Mai and Beroza (2002) calculated autocorrelation of slip distribution and showed that they are explained by von Karman autocorrelation function or functions with power-law dependence at high frequencies. Such scaling about the complexity of earthquake is useful to constrain input source models for strong-motion prediction. However, it should be noted that slip inversion includes many assumptions and smoothing constraints for regularization that greatly affect high-frequency components of seismic waves. Moreover, although previous studies dealt with only final slip distribution, time history of slip will be also important for such purpose in future studies.

Slip models have been obtained for moderate earthquakes of M3–4 (Courboulex et al., 1996; Ide, 2001). Although the resolution is worse than that of large events like the events summarized in **Table 1**, these models can be used to compare slip area and aftershock distribution and to distinguish a simple rupture process from a sequence of discrete subevents. So far, the slip models of the smallest size were determined for five microearthquakes in South African gold mine (Yamada et al., 2005). **Figure 8** compares one of these models (Mw 1.4) with the slip models of the Kobe earthquake (Ide and Takeo, 1997). It is interesting that even though there is a factor of about 1000 changes in the scale, geometric similarities exist between events that differ in 5.4 magnitude unit. Although the resolution is poor, this small event certainly consists of subevents that are also identified directly in observed waves. Some small earthquakes do have complex rupture processes comparable to large events. However, it is still unclear how general such complexity is (see **Chapter 4.18**).

Seismic waves from an earthquake generally have so-called initial phases before the large main phase (e.g., Abercrombie and Mori, 1994; Ellsworth and Beroza, 1995; Iio, 1992, 1995; Umeda, 1990). Sometimes, initial phases are quite different from the main phase. Nakayama and Takeo (1997) determined a slip model for the 1994 Sanriku-Haruka-Oki earthquake (Mw 7.7). This earthquake began with slow propagation at about 2.0 km s^{-1} with small slip during the first 20 s, and later, it was accelerated up to 3.0 km s^{-1} and radiated large seismic waves. The construction of slip models only for initial phases is possible. Shibazaki et al. (2002) determined a slip model for the initial 0.6 s of the Kobe earthquake whose total duration is about 10 s. Comparing with another slip model of an M4 aftershock at similar scale, they found that the rupture propagation of the mainshock is slightly slower than that of the aftershock. Uchide and Ide (2007) developed an inversion method to investigate the initial part of seismic rupture, as well as the whole rupture process, using renormalization of slip distribution. The application to earthquakes with different sizes in the Parkfield, California, suggested the self-similarity of complex earthquake rupture in the wide-scale range (Uchide and Ide, 2010).

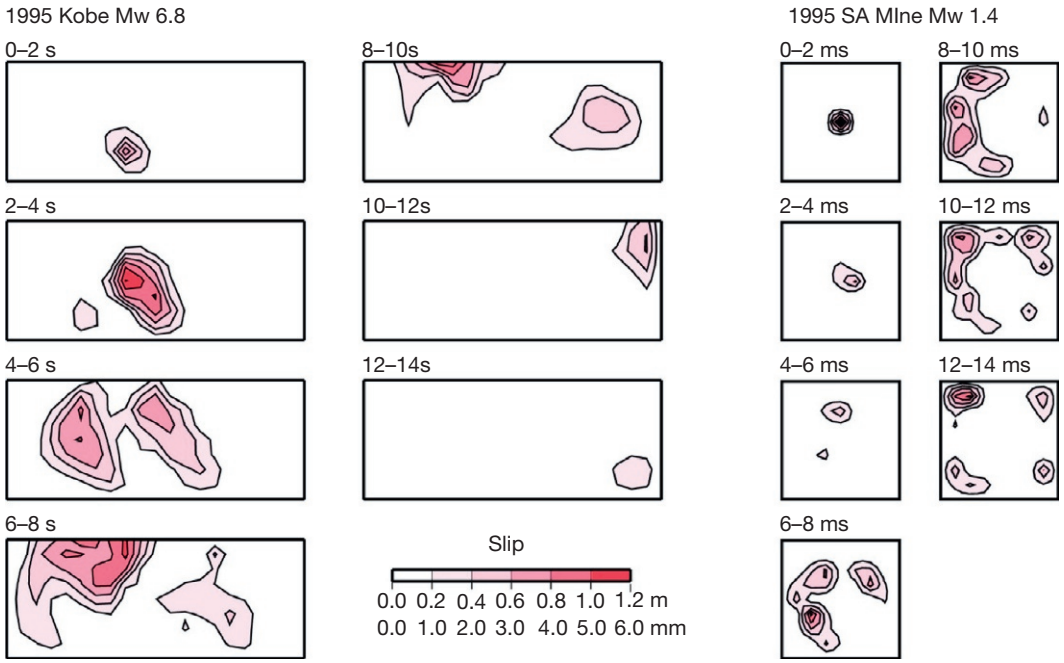


Figure 8 Comparison between two slip models: The 1995 Kobe earthquake and a small earthquake in South Africa gold mine. Slip distributions during each period are shown. The maximum slips are about 1.2 m and 6.0 mm for the Kobe and South African earthquakes, respectively.

4.09.6 Discussion and Conclusion

This chapter reviews the history, formulation, application, and extension of slip inversion using seismic waveform and/or geodetic data. We introduce some popular forms of slip inversion, dividing the problem into basic elements of data, model parameters, and synthetic data. The history of slip inversion started from the early 1980s, after the pioneering study of Trifunac (1974). Although the speed of development has not been rapid since the establishment of the basic forms at around 1990, there are some recent technical improvements, such as the application of various nonlinear minimization methods, objective weighting for regularization constraints, and addition of new kinds of data. The development of back-projection methods provided us the images of rupture process from a different point of view. Except for these improvements, slip inversion is a mature analysis method of seismic sources. Some computer programs are distributed via the Internet (e.g., Kikuchi and Kanamori, 2004). The determination of slip model for relatively large earthquakes using far-field body waves and readily available programs is a kind of routine work or a good exercise for students. Routine reports of such slip models are usually published for significant earthquakes in electronic formats within one or a few days.

It is often overlooked that slip models have only limited reliability. Usually, only gross features such as the location and timing of large slip and average direction and velocity of rupture propagation are well resolved. Even such features can be difficult to identify for some events. An example is the 1999 İzmit earthquake, for which the strong-motion records are not sufficient and their timing information was not correct (Beresnev, 2003; Ide et al., 2005). A project to validate the source inversion using blind tests has implied the fragility of analysis (Mai et al., 2007). Each paper presenting the slip model does not always mention the assumptions and limits of the inversion method, because these problems have been usually investigated in the previous papers. This ambiguity may cause artifacts when we evaluate characteristics and scaling of slip models or derive physical implications from slip models. So far, there is no established way to quantitatively evaluate the reliability of slip models and we have to assess the limitation of models case by case.

The increasing power of computers has contributed to the improvement of slip inversion, increasing the number of data and model parameters. However, the number of model parameters has not increased during this decade. Rather, computer power is recently used to solve nonlinear problems with a large number of unknown parameters using Monte Carlo-like iterative methods. There are large rooms to improve optimization methods for various problems of slip inversion. The degree of freedom is high and the review of this field in this chapter (**Section 4.09.3.1**) is not complete. Another component that requires powerful computers is calculation of Green's function for the heterogeneous Earth structure. One landmark was given by Tsuboi et al. (2003) in the forward simulation of the Denali Fault earthquake using the Earth simulator. Such calculation will increase the resolution of slip inversion not only for a global problem but also for regional earthquakes.

Needless to say, the quality of slip inversion fully depends on the data. Geodetic data will be important for slip inversion to fix the lower-frequency image of the source. InSAR provides unprecedented spatial resolution and the frequency band of GPS partly overlaps with that of broadband seismometers. Seamless investigation for pre-, co-, and postseismic processes has become feasible using high-sampling GPS data (e.g., Yokota, 2013). Nevertheless, seismic data, especially near-field data, are essential to resolve temporal change in detail. Although the inland earthquakes can be resolved quite well, they are just a small fraction of all earthquakes in the world. Most events occur beneath the ocean, where little local data are available for slip inversion. The 2003 Tokachi-Oki, Japan, earthquake (Mw 8.3; Hirata et al., 2003) and the 2011 Tohoku-Oki earthquake were rare examples recorded by a number of near-field strong-motion stations. The records of cabled ocean bottom stations are quite expensive, but to increase such instruments is an important task for the study of seismic source.

There have been many slip models for more than 100 earthquakes, providing us with the images of fault slip and information about the physics of earthquakes. However, there are only a small number of earthquakes whose rupture processes were resolved well enough for the discussion of fault dynamics. The Chi-Chi, Landers, Northridge, Kobe, Loma Prieta, Imperial Valley, and Tottori (**Table 1**) are those events. Recent very large earthquakes such as the Sumatra, Maule, and Tohoku-Oki are another group of events, which are well resolved even by teleseismic waves. We need more case studies together with quality controls of the models. It is also important to develop analysis methods that enable broadband slip inversion that can use high-frequency waves in complex structures.

Acknowledgments

The author thanks Professor M. Hashimoto for the fault data of Figure 1 and Professor M. Matsu'ura and Dr. H. Aochi for their useful suggestions. The fault models compiled by Drs. M. Mai and D. Wald were quite helpful for comparison between models. The catalog of slip models for Japanese earthquakes compiled by Professors R. Sato and K. Koketsu was also useful. The digital data of the models for the Tohoku-Oki earthquake are provided by Drs. W. Suzuki, Y. Yokota, and Y. Yagi. The manuscript was greatly improved by comments from Professor M. Ishii. Some figures are prepared using the generic mapping tool (Wessel and Smith, 1991). This work is partly supported by Grant-in-Aid for Scientific Research and DaiDaiToku Project, Ministry of Education, Sports, Science and Technology, Japan.

References

Abercrombie R and Mori J (1994) Local observations of the onset of a large earthquake: 28 June 1992 Landers, California. *Bulletin of the Seismological Society of America* 84: 725–734.

Akaike H (1980) Likelihood and Bayes procedure. In: Bernardo JM, DeGroot MH, Lindley DV, and Smith AFM (eds.) *Bayesian Statistics*, pp. 143–166. Valencia, Spain: University Press.

Aki K (1967) Scaling law of seismic spectrum. *Journal of Geophysical Research* 72: 1217–1231.

Aki K (1968) Seismic displacements near a fault. *Journal of Geophysical Research* 73: 5359–5376.

Aki K and Richards PG (2002) *Quantitative Seismology*, 2nd edn. Sausalito, CA: University Science Book.

Anderson JG (2003) Strong-motion seismology. In: Lee WHK, Kanamori H, Jennings PC, and Kisslinger C (eds.) *Earthquake and Engineering Seismology, Part B*, pp. 937–965. London: Academic Press.

Andrews DJ (1976) Rupture velocity of plane strain shear cracks. *Journal of Geophysical Research* 81: 5679–5687.

Aochi H and Fukuyama E (2002) Three-dimensional nonplanar simulation of the 1992 Landers earthquake. *Journal of Geophysical Research* 107. http://dx.doi.org/10.1029/2000JB000061.

Aochi H, Madariaga R, and Fukuyama E (2003) Constraint of fault parameters inferred from nonplanar fault modeling. *Geochemistry, Geophysics, Geosystems* 4. http://dx.doi.org/10.1029/2001GC000207.

Aoki Y, Segall P, Kato T, Cervelli P, and Shimada S (1999) Imaging magma transport during the 1997 seismic swarm off the Izu Peninsula, Japan. *Science* 286: 927–930.

Archuleta RJ (1984) A faulting model for the 1979 Imperial Valley earthquake. *Journal of Geophysical Research* 89: 4559–4585.

Aster RC, Borchers B, and Thurber CH (2012) *Parameter Estimation and Inverse Problems*, 2nd edn. Oxford, UK: Academic Press.

Benioff H, Press F, and Smith S (1961) Excitation of the free oscillations of earth by earthquakes. *Journal of Geophysical Research* 66: 605–619.

Beresnev IA (2003) Uncertainties in finite-fault slip inversions: To what extent to believe? (A critical review). *Bulletin of the Seismological Society of America* 93: 2445–2458.

Bernard P and Madariaga R (1984) A new asymptotic method for the modeling of near-field accelerograms. *Bulletin of the Seismological Society of America* 74: 539–557.

Beroza GC (1991) Near-source modeling of the Loma Prieta earthquake: Evidence for heterogeneous slip and implications for earthquake hazard. *Bulletin of the Seismological Society of America* 81: 1603–1621.

Beroza GC (1995) Seismic source modeling. *Reviews of Geophysics* 33(S1): 299–308.

Beroza GC and Mikumo T (1996) Short slip duration in dynamic rupture in the presence of heterogeneous fault properties. *Journal of Geophysical Research* 101: 22449–22460.

Beroza GC and Spudich P (1988) Linearized inversion for fault rupture behavior: Application to the 1984 Morgan Hill, California, earthquake. *Journal of Geophysical Research* 93: 6275–6296.

Boatwright J and Choy GL (1986) Teleseismic estimates of the energy radiated by shallow earthquakes. *Journal of Geophysical Research* 91: 2095–2112.

Bouchon M (1976) Teleseismic body wave radiation from a seismic source in a layered medium. *Geophysical Journal of the Royal Astronomical Society* 47: 515–530.

Bouchon M (1981) A simple method to calculate Green's functions for elastic layered media. *Bulletin of the Seismological Society of America* 71: 959–971.

Bouchon M (1997) The state of stress on some faults on the San Andreas system as inferred from near-field strong motion data. *Journal of Geophysical Research* 102: 11731–11744.

Bouchon M, Toksöz MN, Karabulut H, et al. (2002) Space and time evolution of rupture and faulting during the 1999 İzmit (Turkey) earthquake. *Bulletin of the Seismological Society of America* 92: 256–266.

Brune JN (1970) Tectonic stress and the spectra of seismic shear waves from earthquakes. *Journal of Geophysical Research* 75: 4997–5009.

Chi WC, Dreger D, and Kaverina A (2001) Finite-source modeling of the 1999 Taiwan (Chi-Chi) earthquake derived from a dense strong-motion network. *Bulletin of the Seismological Society of America* 91: 1144–1157.

Cho I and Nakanishi I (2000) Investigation of the three-dimensional fault geometry ruptured by the 1995 Hyogo-Ken Nanbu earthquake using strong-motion and geodetic data. *Bulletin of the Seismological Society of America* 90: 450–467.

Cocco M, Spudich P, and Tinti E (2006) On the mechanical work absorbed on faults during earthquake ruptures. In: Abercrombie R, et al. (ed.) *Geophysical Monograph Series 170, Earthquakes: Radiated Energy and the Physics of Faulting*, pp. 237–254. Washington, DC: American Geophysical Union.

Cochard A and Madariaga R (1994) Dynamic faulting under rate-dependent friction. *Pure and Applied Geophysics* 142: 419–445.

Cohee BP and Beroza GC (1994) Slip distribution of the 1992 Landers earthquake and its implications for earthquake source mechanics. *Bulletin of the Seismological Society of America* 84: 692–712.

Cotton F and Campillo M (1995) Frequency domain inversion of strong motions: Application to the 1992 Landers earthquake. *Journal of Geophysical Research* 100: 3961–3975.

Courboulex F, Virieux J, Deschamps A, Gilbert D, and Zollo A (1996) Source investigation of a small event using empirical Green's functions and simulated annealing. *Geophysical Journal International* 125: 768–780.

Das S and Aki K (1977) A numerical study of two-dimensional spontaneous rupture propagation. *Geophysical Journal of the Royal Astronomical Society* 50: 643–668.

Das S and Henry C (2003) Spatial relation between main earthquake slip and its aftershock distribution. *Reviews of Geophysics* 41. http://dx.doi.org/10.1029/2002RG000119.

Das S and Kostrov BV (1990) Inversion for seismic slip rate history and distribution with stabilizing constraints: Application to the 1986 Andreanof Islands earthquake. *Journal of Geophysical Research* 95: 6899–6913.

Das S and Scholz CH (1983) Why large earthquakes do not nucleate at shallow depths. *Nature* 305: 621–623.

Day SM (1982a) Three-dimensional simulation of spontaneous rupture: The effect of nonuniform prestress. *Bulletin of the Seismological Society of America* 72: 1881–1902.

Day SM (1982b) Three dimensional finite differential simulation of fault dynamics: Rectangular faults with fixed rupture velocity. *Bulletin of the Seismological Society of America* 72: 705–727.

Delouis B, Giardini D, Lundgren P, and Salichon J (2002) Joint inversion of InSAR, GPS, teleseismic, and strong-motion data for the spatial and temporal distribution of earthquake slip: Application to the 1999 İzmit Mainshock. *Bulletin of the Seismological Society of America* 92: 278–299.

Dziewonski AM and Woodhouse JH (1983) An experiment in systematic study of global seismicity: Centroid-moment tensor solutions for 201 moderate and large earthquakes of 1981. *Journal of Geophysical Research* 88: 3247–3271.

Ellsworth WL and Beroza GC (1995) Seismic evidence for an earthquake nucleation phase. *Science* 268: 851–855.

Fletcher JB and Spudich P (1998) Rupture characteristics of the three M 4.7 (1992–1994) Parkfield earthquakes. *Journal of Geophysical Research* 103: 835–854.

Frankel A (2004) Rupture process of the M 7.9 Denali fault, Alaska, earthquake: Subevents, directivity, and scaling of high-frequency ground motions. *Bulletin of the Seismological Society of America* 94: S234–S255.

Fuchs K and Müller G (1971) Computation of synthetic seismograms with the reflectivity method and comparison with observations. *Geophysical Journal of the Royal Astronomical Society* 23: 417–433.

Fujii Y, Satake K, Sakai S, Shinohara M, and Kanazawa T (2011) Tsunami source of the 2011 off the Pacific coast of Tohoku Earthquake. *Earth, Planets and Space* 63: 815–820.

Fukahata Y, Nishitani A, and Matsu'ura M (2004) Geodetic data inversion using ABIC to estimate slip history during one earthquake cycle with viscoelastic slip-response functions. *Geophysical Journal International* 156: 140–153.

Fukuda J and Johnson M (2008) *Bulletin of the Seismological Society of America* 98: 1128–1146.

Fukuda J and Johnson M (2010) Mixed linear–non-linear inversion of crustal deformation data: Bayesian inference of model, weighting and regularization parameters. *Geophysical Journal International* 181: 1441–1458.

Fukuyama E and Irikura K (1986) Rupture process of the 1983 Japan Sea (Akita-Oki) earthquake using a waveform inversion method. *Bulletin of the Seismological Society of America* 76: 1623–1640.

Fukuyama E and Mikumo T (1993) Dynamic rupture analysis: Inversion for the source process of the 1990 Izu-Oshima, Japan, earthquake (M¼6.5). *Journal of Geophysical Research* 98: 6529–6542.

Fukuyama E, Mikumo T, and Olsen KB (2003) Estimation of the critical slip-weakening distance: Theoretical background. *Bulletin of the Seismological Society of America* 93: 1835–1840.

Geller RJ (1976) Scaling relations for earthquake source parameters and magnitudes. *Bulletin of the Seismological Society of America* 66: 1501–1523.

Graves RW and Wald DJ (2001) Resolution analysis of finite fault source inversion using one- and three-dimensional Green's functions. Part 1: Strong motions. *Journal of Geophysical Research* 106: 8745–8766.

Guatteri M and Spudich P (1998) Coseismic temporal changes of slip direction; the effect of absolute stress on dynamic rupture. *Bulletin of the Seismological Society of America* 88: 777–789.

Guatteri M and Spudich P (2000) What can strong-motion data tell us about slip-weakening fault-friction laws. *Bulletin of the Seismological Society of America* 90: 98–116.

Guatteri M, Spudich P, and Beroza GC (2001) Inferring rate and state friction parameters from a rupture model of the 1995 Hyogo-ken Nanbu (Kobe) Japan earthquake. *Journal of Geophysical Research* 106: 26511–26521.

Gülen L, Pinar A, Kalafat D, et al. (2002) Surface fault breaks, aftershock distribution and rupture process of the 17 August 1999 İzmit, Turkey, earthquake. *Bulletin of the Seismological Society of America* 92: 230–244.

Gutenberg B and Richter CF (1944) Frequency of earthquakes in California. *Bulletin of the Seismological Society of America* 34: 185–188.

Hartzell SH (1978) Earthquake aftershocks as Green's functions. *Geophysical Research Letters* 5: 1–4.

Hartzell S (1989) Comparison of seismic waveform inversion results for the rupture history of a finite fault: Application to the 1986 North Palm Springs, California, earthquake. *Journal of Geophysical Research* 94: 7515–7534.

Hartzell SH and Heaton TH (1983) Inversion of strong ground motion and teleseismic waveform data for the fault rupture history of the 1979 Imperial Valley, California, earthquake. *Bulletin of the Seismological Society of America* 73: 1553–1583.

Hartzell S and Langer C (1993) Importance of model parameterization in finite fault inversions: Application to the 1974 Mw 8.0 Peru earthquake. *Journal of Geophysical Research* 98: 22123–22134.

Hartzell S and Liu P (1995) Determination of earthquake source parameters using a hybrid global search algorithm. *Bulletin of the Seismological Society of America* 85: 516–524.

Hartzell S, Liu P, and Mendoza C (1996) The 1994 Northridge, California, earthquake: Investigation of rupture velocity, risetime, and high-frequency radiation. *Journal of Geophysical Research* 101: 20091–20108.

Hartzell SH, Stewart GS, and Mendoza C (1991) Comparison of L1 and L2 norms in a teleseismic waveform inversion for the slip history of the Loma Prieta, California, earthquake. *Bulletin of the Seismological Society of America* 81: 1518–1539.

Hashimoto M (2005) How close to the Kobe Earthquake did seismology come? In: *Proceedings of the 2005 Joint Meeting for Earth and Planetary Science, Chiba, Japan,* S094-001.

Haskell NA (1964) Total energy and spectral density of elastic wave radiation from propagating faults. *Bulletin of the Seismological Society of America* 54: 1811–1841.

Haskell NA (1969) Elastic displacements in the near-field of a propagating fault. *Bulletin of the Seismological Society of America* 59: 865–908.

Heaton TH (1990) Evidence for and implications of self-healing pulses of slip in earthquake rupture. *Physics of the Earth and Planetary Interiors* 64: 1–20.

Hellweg M and Boatwright J (1999) Mapping the rupture process of moderate earthquakes by inverting accelerograms. *Journal of Geophysical Research* 104: 7319–7328.

Hernandez B, Cotton F, and Campillo M (1999) Contribution of radar interferometry to a two-step inversion of the kinematic process of the 1992 Landers earthquake. *Journal of Geophysical Research* 104: 13083–13099.

Hirata K, Baba T, Sugioka H, et al. (2003) Ocean-bottom near-field observation to the 2003 Tokachi-oki earthquake with cabled seafloor observatory. *Earth Observing System Transactions, American Geophysical Union* 84: S52L-04.

Horikawa H, Hirahara K, Umeda Y, Hashimoto M, and Kusano F (1996) Simultaneous inversion of geodetic and strong-motion data for the source process of the Hyogo-ken Nanbu, Japan, earthquake. *Journal of the Physics of the Earth* 44: 455–471.

Ichinose GA, Somerville P, Thio HK, Matsushima S, and Sato T (2005) Rupture process of the 1948 Fukui earthquake (M 7.1) from the joint inversion of seismic waveform and geodetic data. *Journal of Geophysical Research* 110. http://dx.doi.org/10.1029/2004JB003437.

Ichinose GA, Thio HK, Somerville PG, Sato T, and Ishii T (2003) Rupture process of the 1944 Tonankai earthquake (Ms 8.1) from the inversion of teleseismic and regional seismograms. *Journal of Geophysical Research* 108. http://dx.doi.org/10.1029/2003JB002393.

Ida Y (1972) Cohesive force across the tip of a longitudinal-shear crack and Griffith's specific surface energy. *Journal of Geophysical Research* 77: 3796–3805.

Ide S (1999) Source process of the 1997 Yamaguchi, Japan, earthquake analyzed in different frequency bands. *Geophysical Research Letters* 26: 1973–1976.

Ide S (2001) Complex source processes and the interaction of moderate earthquakes during the earthquake swarm in the Hida-Mountains, Japan, 1998. *Tectonophysics* 334: 35–54.

Ide S (2002) Estimation of radiated energy of finite-source earthquake models. *Bulletin of the Seismological Society of America* 92: 2294–3005.

Ide S (2003) Fracture surface energy of natural earthquakes from the viewpoint of seismic observations. *Bulletin of the Earthquake Research Institute* 78: 59–65.

Ide S and Aochi H (2005) Earthquake as multi-scale dynamic ruptures with heterogeneous fracture surface energy. *Journal of Geophysical Research* 110. http://dx.doi.org/10.1029/2004JB003591.

Ide S, Baltay A, and Beroza GC (2011) Shallow dynamic overshoot and energetic deep rupture in the 2011 Mw 9.0 Tohoku-Oki earthquake. *Science* 332: 1426–1429.

Ide S and Beroza GC (2001) Does apparent stress vary with earthquake size? *Geophysical Research Letters* 28: 3349–3352.

Ide S, Beroza GC, and McGuire JJ (2005) Imaging earthquake source complexity. In: Levander A and Nolet G (eds.) *Geophysical Monograph Series 157, Seismic Earth: Array Analysis of Broadband Seismograms*, pp. 117–135. Washington, DC: American Geophysical Union.

Ide S and Takeo M (1996) The dynamic rupture process of the 1993 Kushiro-oki earthquake. *Journal of Geophysical Research* 101: 5661–5675.

Ide S and Takeo M (1997) Determination of constitutive relations of fault slip based on seismic wave analysis. *Journal of Geophysical Research* 102: 27379–27391.

Ihmlé PF (1998) On the interpretation of subevents in teleseismic waveforms: The 1994 Bolivia deep earthquake revisited. *Journal of Geophysical Research* 103: 17919–17932.

Iinuma T, Hino R, Kido M, et al. (2012) Coseismic slip distribution of the 2011 off the Pacific Coast of Tohoku Earthquake (M9.0) refined by means of seafloor geodetic data. *Journal of Geophysical Research* 117: B07409.

Iio Y (1992) Slow initial phase of the P-wave velocity pulse generated by microearthquakes. *Geophysical Research Letters* 19: 477–480.

Iio Y (1995) Observations of the slow initial phase generated by microearthquakes—Implications for earthquake nucleation and propagation. *Journal of Geophysical Research* 100: 15333–15349.

Imanishi K and Takeo M (2002) An inversion method to analyze rupture processes of small earthquakes using stopping phases. *Journal of Geophysical Research* 107. http://dx.doi.org/10.1029/2001JB000201.

Imanishi K, Takeo M, Ellsworth WL, et al. (2004) Source parameters and rupture velocities of microearthquakes in western Nagano, Japan, determined using stopping phases. *Bulletin of the Seismological Society of America* 94: 1762–1780.

Ishii M (2011) High-frequency rupture properties of the Mw 9.0 off the Pacific coast of Tohoku Earthquake. *Earth, Planets and Space* 63: 609–614.

Ishii M, Shearer PM, Houston H, and Vidale JE (2005) Extent, duration and speed of the 2004 Sumatra–Andaman earthquake imaged by the Hi-Net array. *Nature* 435: 933–936.

Iwata T (1991) Details of strong ground motions near earthquake foci and studies of source processes using these records. *Zisin* 2(44): 315–327 (in Japanese with English abstract).

Iwata T and Sekiguchi H (2002) Source model of the 2000 Tottori-ken Seibu earthquake and near-source strong ground motion. In: *Proceedings of the 11th Japan Earthquake Engineering Symposium*, pp. 125–128 (in Japanese with English abstract).

Jackson DD (1979) The use of a priori data to resolve non-uniqueness in linear inversion. *Geophysical Journal of the Royal Astronomical Society* 57: 137–157.

Jackson DD and Matsu'ura M (1985) A Bayesian approach to nonlinear inversion. *Journal of Geophysical Research* 90: 581–591.

Ji C, Helmberger DV, Wald DJ, and Ma K-F (2003) Slip history and dynamic implications of the 1999 Chi-Chi, Taiwan earthquake. *Journal of Geophysical Research* 108. http://dx.doi.org/10.1029/2002JB001764.

Ji C, Wald DJ, and Helmberger DV (2002a) Source description of the 1999 Hector Mine, California, earthquake. Part I: Wavelet domain inversion theory and resolution analysis. *Bulletin of the Seismological Society of America* 92: 1192–1207.

Ji C, Wald DJ, and Helmberger DV (2002b) Source description of the 1999 Hector Mine, California, earthquake. Part II: Complexity of slip history. *Bulletin of the Seismological Society of America* 92: 1208–1226.

Kakehi Y and Irikura K (1996) Estimation of high-frequency wave radiation areas on the fault plane by the envelope inversion of acceleration seismograms. *Geophysical Journal International* 125: 892–900.

Kakehi Y and Irikura K (1997) High-frequency radiation process during earthquake faulting—Envelope inversion of acceleration seismograms from the 1993 Hokkaido-Nansei-Oki, Japan, earthquake. *Bulletin of the Seismological Society of America* 87: 904–917.

Kakehi Y, Irikura K, and Hoshiba M (1996) Estimation of high-frequency wave radiation areas on the fault plane of the 1995 Hyogo-ken Nanbu earthquake by the envelope inversion of acceleration seismograms. *Journal of the Physics of the Earth* 44: 505–517.

Kanamori H (2004) The diversity of the physics of earthquakes. *Proceedings of the Japan Academy Series B* 80: 297–316.

Kanamori H and Anderson DL (1975) Theoretical basis of some empirical relations in seismology. *Bulletin of the Seismological Society of America* 65: 1073–1095.

Kanamori H and Brodsky EE (2001) The physics of earthquakes. *Physics Today* 54: 34–39.

Kanamori H and Stewart GS (1978) Seismological aspects of the Guatemala earthquake of February 4, 1976. *Journal of Geophysical Research* 83: 3427–3434.

Kennett BLN (2001) *The Seismic Wavefield Volume I: Introduction and Theoretical Development*. Cambridge, UK: Cambridge University Press.

Kennett BLN and Kerry NJ (1979) Seismic waves in a stratified half space. *Geophysical Journal of the Royal Astronomical Society* 57: 557–583.

Kikuchi M (1991) Complexity of earthquake source processes. *Zisin* 2(44): 301–314 (in Japanese with English abstract).

Kikuchi M and Kanamori H (1982) Inversion of complex body waves. *Bulletin of the Seismological Society of America* 72: 491–506.

Kikuchi M and Kanamori H (1986) Inversion of complex body waves – II. *Physics of the Earth and Planetary Interiors* 43: 205–222.
Kikuchi M and Kanamori H (1991) Inversion of complex body waves – III. *Bulletin of the Seismological Society of America* 81: 2335–2350.
Kikuchi M and Kanamori H (1994) The mechanism of the deep Bolivia earthquake of June 9, 1994. *Geophysical Research Letters* 21: 2341–2344.
Kikuchi M and Kanamori H (2004) Notes on teleseismic body-wave inversion program. http://www.eri.u-tokyo.ac.jp/ETAL/KIKUCHI (accessed December 2012).
Kikuchi M, Nakamura M, Yamada M, et al. (1999) Source parameters of the 1948 Fukui earthquake inferred from low-gain strong-motion records. *Zisin* 2(52): 121–128 (in Japanese with English abstract).
Kikuchi M, Nakamura M, and Yoshikawa K (2003) Source rupture process of the 1944 Tonankai earthquake and the 1945 Mikawa earthquake derived from low-gain seismograms. *Earth, Planets and Space* 55: 159–172.
Kind R (1976) Computation of reflection coefficients for layered media. *Journal of Geophysics* 42: 191–200.
Kiser E, Ishii M, Langmuir CH, Shearer PM, and Hirose H (2011) Insights into the mechanism of intermediate-depth earthquakes from source properties as imaged by back projection of multiple seismic phases. *Journal of Geophysical Research* 116: B06310. http://dx.doi.org/10.1029/2010JB007831.
Kohketsu K (1985) The extended reflectivity method for synthetic near-field seismograms. *Journal of the Physics of the Earth* 33: 121–131.
Koketsu K, Yokota Y, Nishimura N, et al. (2011) A unified source model for the 2011 Tohoku earthquake. *Earth and Planetary Science Letters* 310: 480–487.
Koketsu K, Yoshida S, and Higashihara H (1998) A fault model of the 1995 Kobe earthquake derived from the GPS data on the Akashi Kaikyo Bridge and other datasets. *Earth, Planets and Space* 50: 803–811.
Koper KD, Hutko AR, Lay T, Ammon CJ, and Kanamori H (2011) Frequency-dependent rupture process of the 2011 Mw 9.0 Tohoku Earthquake: Comparison of short-period P wave backprojection images and broadband seismic rupture models. *Earth, Planets and Space* 63: 599–602.
Kostrov VV (1974) Seismic moment and energy of earthquakes, and seismic flow of rock. *Izvestiya Earth Physics* 1: 23–40.
Krüger F and Ohrnberger M (2005) Tracking the rupture of the Mw¼9.3 Sumatra earthquake over 1,150 km at teleseismic distance. *Nature* 435: 937–939.
Langston CA and Helmberger DV (1975) A procedure for modeling shallow dislocation sources. *Geophysical Journal of the Royal Astronomical Society* 42: 117–130.
Lawson CL and Hanson RJ (1974) *Solving Least Squares Problems*. Englewood Cliffs, NJ: Prentice-Hall.
Lee WHK (2002) Challenges in observational seismology. In: Lee WHK, Kanamori H, Jennings PC, and Kisslinger C (eds.) *Earthquake and Engineering Seismology*, pp. 269–281. London: Academic Press.
Lee WHK, Shin TC, Kuo KW, Chen KC, and Wu CF (2001) CWB free-field strong-motion data from the 21 September Chi-Chi, Taiwan, earthquake. *Bulletin of the Seismological Society of America* 91: 1370–1376.
Liu P and Archuleta RJ (2004) A new nonlinear finite fault inversion with three-dimensional Green's functions: Application to the 1989 Loma Prieta, California, earthquake. *Journal of Geophysical Research* 109. http://dx.doi.org/10.1029/2003JB002625.
Ma K-F, Mori J, Lee S-J, and Yu SB (2001) Spatial and temporal distribution of slip for the 1999 Chi-Chi, Taiwan, earthquake. *Bulletin of the Seismological Society of America* 91: 1069–1087.
Madariaga R (1977) High-frequency radiation from crack (stress drop) models of earthquake faulting. *Geophysical Journal of the Royal Astronomical Society* 51: 625–651.
Mai PM and Beroza GC (2000) Source scaling properties from finite-fault-rupture models. *Bulletin of the Seismological Society of America* 90: 604–614.
Mai PM and Beroza GC (2002) A spatial random field model to characterize complexity in earthquake slip. *Journal of Geophysical Research* 107. http://dx.doi.org/10.1029/2001JB000588.
Mai PM, Burjanek J, Delouis B, et al. (2007) Earthquake source inversion blind test: Initial results and further developments. *Earth Observing System Transactions, American Geophysical Union* 88: S53C-08.
Mai PM, Spudich P, and Boatwright J (2005) Hypocenter locations in finite-source rupture models. *Bulletin of the Seismological Society of America* 95: 965–980.
Matsu'ura M (1991) Development of inversion theory in geophysics. *Zisin* 2(44): 53–62 (in Japanese with English abstract).
Matsu'ura M and Sato T (1997) Loading mechanism and scaling relations of large interplate earthquakes. *Tectonophysics* 277: 189–198.
McGuire JJ, Zhao L, and Jordan TH (2001) Teleseismic inversion for the second-degree moments of earthquake space-time distributions. *Geophysical Journal International* 145: 661–678.
McGuire JJ, Zhao L, and Jordan TH (2002) Predominance of unilateral rupture for a global catalog of large earthquakes. *Bulletin of the Seismological Society of America* 92: 3309–3317.
Mendez AJ and Anderson JG (1991) The temporal and spatial evolution of the 19 September 1985 Michoacan earthquake as inferred from near-source ground-motion records. *Bulletin of the Seismological Society of America* 81: 844–861.
Mendoza C and Hartzell SH (1988a) Inversion for slip distribution using teleseismic P waveforms: North Palm Springs, Borah Peak, and Michoacan earthquakes. *Bulletin of the Seismological Society of America* 78: 1092–1111.
Mendoza C and Hartzell SH (1988b) Aftershock patterns and main shock faulting. *Bulletin of the Seismological Society of America* 78: 1438–1449.
Meng L, Inbal A, and Ampuero JP (2011) A window into the complexity of the dynamic rupture of the 2011 Mw 9 Tohoku-Oki earthquake. *Geophysical Research Letters* 38: L00G07. http://dx.doi.org/10.1029/2011GL048118.
Menke W (2012) *Geophysical Data Analysis: Discrete Inverse Theory*, MATLAB edn. San Diego, CA: Academic Press.
Mikumo T, Olsen KB, Fukuyama E, and Yagi Y (2003) Stress-breakdown time and slip-weakening distance inferred from slip-velocity functions on earthquake faults. *Bulletin of the Seismological Society of America* 93: 264–282.
Miyatake T (1992) Dynamic rupture processes of inland earthquakes in Japan: Weak and strong asperities. *Geophysical Research Letters* 19: 1041–1044.
Miyazaki S, Larson KM, Choi K, et al. (2004) Modeling the rupture process of the 2003 September 25 Tokachi-Oki (Hokkaido) earthquake using 1-Hz GPS data. *Geophysical Research Letters* 31: L21603. http://dx.doi.org/10.1029/2004GL021457.
Miyazaki S, McGuire JJ, and Segall P (2003) A transient subduction zone slip episode in southwest Japan observed by the nationwide GPS array. *Journal of Geophysical Research* 108. http://dx.doi.org/10.1029/2001JB000456.
Miyazaki S, Segall P, McGuire JJ, Kato T, and Hatanaka Y (2006) Spatial and temporal evolution of stress and slip rate during the 2000 Tokai slow earthquake. *Journal of Geophysical Research* 11. http://dx.doi.org/10.1029/2004JB003426.
Mori J and Hartzell S (1990) Source inversion of the 1988 Upland, California, earthquake: Determination of a fault plane for a small event. *Bulletin of the Seismological Society of America* 80: 507–518.
Nakahara H, Nishimura T, Sato H, and Ohtake M (1998) Seismogram envelop inversion for the spatial distribution of high-frequency energy radiation from the earthquake fault: Application to the 1994 far east off Sanriku earthquake, Japan. *Journal of Geophysical Research* 103: 855–867.
Nakahara H, Sato H, Ohtake M, and Nishimura T (1999) Spatial distribution of high-frequency energy radiation on the fault of the 1995 Hyogo-Ken Nanbu, Japan, earthquake (Mw 6.9) on the basis of the seismogram envelope inversion. *Bulletin of the Seismological Society of America* 89: 22–35.
Nakayama W and Takeo M (1997) Slip history of the 1994 Sanriku-Haruka-Oki, Japan, earthquake deduced from strong-motion data. *Bulletin of the Seismological Society of America* 87: 918–931.
Oglesby DD, Archuleta RJ, and Nielsen SB (1998) Earthquakes on dipping faults: The effects of broken symmetry. *Science* 280: 1055–1059.
Oglesby DD and Day SM (2001) Fault geometry and the dynamics of the 1999 Chi-Chi (Taiwan) earthquake. *Bulletin of the Seismological Society of America* 91: 1099–1111.
Okada Y (1985) Surface deformation due to shear and tensile faults in a half-space. *Bulletin of the Seismological Society of America* 75: 1135–1154.
Okada T and Hasegawa A (2003) The M7.1 May 26, 2003 offshore Miyagi Prefecture earthquake in northeast Japan: Source process and aftershock distribution of an intra-slab event. *Earth, Planets and Space* 55: 731–739.
Okada T, Umino N, Ito Y, et al. (2001) Source processes of 15 September 1998 M 5.0 Sendai, northeastern Japan, earthquake and its M 3.8 foreshock by waveform inversion. *Bulletin of the Seismological Society of America* 91: 1607–1618.
Okamoto T and Miyatake T (1989) Effects of near source seafloor topography on long-period teleseismic P waveforms. *Geophysical Research Letters* 16: 1309–1312.
Oliver J and Murphy L (1971) WWNSS: Seismology's global network of observing stations. *Science* 174: 254–261.
Olsen K, Madariaga R, and Archuleta R (1997) Three-dimensional dynamic simulation of the 1992 Landers earthquake. *Science* 278: 834–838.
Olson AH and Anderson JG (1988) Implications of frequency-domain inversion of earthquake ground motions for resolving the space-time dependence of slip on an extended fault. *Geophysical Journal* 94: 443–455.
Olson AH and Apsel RJ (1982) Finite faults and inverse theory with applications to the 1979 Imperial Valley earthquake. *Bulletin of the Seismological Society of America* 72: 1969–2001.
Ozawa S, Nishimura T, Suito H, et al. (2011) Coseismic and postseismic slip of the 2011 magnitude-9 Tohoku-Oki earthquake. *Nature* 475: 373–376.

Peyrat S, Olsen K, and Madariaga R (2001) Dynamic modeling of the 1992 Landers earthquake. *Journal of Geophysical Research* 106: 26467–26482.

Press F, Ben-Menahem A, and Toksöz MN (1961) Experimental determination of earthquake fault length and rupture velocity. *Journal of Geophysical Research* 66: 3471–3485.

Quin H (1990) Dynamic stress drop and rupture dynamics of the October 15, 1979 Imperial Valley, California, earthquake. *Tectonophysics* 175: 93–117.

Romano F, Piatanesi A, Lorito S, et al. (2012) Clues from joint inversion of tsunami and geodetic data of the 2011 Tohoku-oki earthquake. *Scientific Reports* 2(385): 1–8.

Rost S and Thomas C (2002) Array seismology: Methods and applications. *Reviews of Geophysics* 40: 1008. http://dx.doi.org/10.1029/2000RG000100.

Rothman DH (1986) Automatic estimation of large residual statics corrections. *Geophysics* 51: 332–346.

Ruiz S and Madariaga R (2011) Determination of the friction law parameters of the Mw 6.7 Michilla earthquake in northern Chile by dynamic inversion. *Geophysical Research Letters* 38: L09317.

Saikia CK (1994) Modified frequency-wavenumber algorithm for regional seismograms using Filon's quadrature: Modelling of Lg waves in eastern North America. *Geophysical Journal International* 118: 142–158.

Salichon J, Lundgren P, Delouis B, and Giardini D (2004) Slip history of the 16 October 1999 Mw 7.1 Hector Mine earthquake (California) from the inversion of InSAR, GPS, and Teleseismic Data. *Bulletin of the Seismological Society of America* 94: 2015–2027.

Salichon J, Delouis B, Lundgren P, et al. (2003) Joint inversion of broadband teleseismic and interferometric synthetic aperture radar (InSAR) data for the slip history of the Mw = 7.7, Nazca ridge (Peru) earthquake of 12 November 1996. *Journal of Geophysical Research* 108. http://dx.doi.org/10.1029/2001JB000913.

Satake K (1989) Inversion of tsunami waveforms for the estimation of heterogeneous fault motion of large submarine earthquakes: The 1968 Tokachi-oki and 1983 Japan Sea earthquakes. *Journal of Geophysical Research* 94: 5627–5636.

Satake K (1993) Depth distribution of coseismic slip along the Nankai trough, Japan, from joint inversion of geodetic and tsunami data. *Journal of Geophysical Research* 98: 4553–4565.

Sato H, Nakahara H, and Ohtake M (1997) Synthesis of scattered energy density for nonspherical radiation from a point shear-dislocation source based on the radiative transfer theory. *Physics of the Earth and Planetary Interiors* 104: 1–13.

Scherbaum F (1996) *Of Poles and Zeros: Fundamentals of Digital Seismology*. Dordrecht: Kluwer.

Scholz CH (2002) *The Mechanics of Earthquakes and Faulting*, 2nd edn. Cambridge, UK: Cambridge University Press.

Sekiguchi H, Irikura K, and Iwata T (1998) Detailed source process of the 1995 Hyogo-ken Nanbu (Kobe) earthquake using near-source strong ground motion data. In: *Proceedings of the 10th Japan Earthquake Engineering Symposium,* vol. 1, pp. 67–72.

Sekiguchi H, Irikura K, Iwata T, Kakehi Y, and Hoshiba M (1996) Minute locating of faulting beneath Kobe and the waveform inversion of the source process during the 1995 Hyogo-ken Nanbu, Japan, earthquake using strong ground motion records. *Journal of the Physics of the Earth* 44: 473–487.

Sekiguchi H and Iwata T (2002) Rupture process of the 1999 Kocaeli, Turkey, earthquake estimated from strong-motion waveforms. *Bulletin of the Seismological Society of America* 92: 300–311.

Semmane F, Cotton F, and Campillo M (2005) The 2000 Tottori earthquake: A shallow earthquake with no surface rupture and slip properties controlled by depth. *Journal of Geophysical Research* 110. http://dx.doi.org/10.1029/2004JB003194.

Shao G, Li X, Ji C, and Maeda T (2011) Focal mechanism and slip history of the 2011 Mw 9.1 off the Pacific coast of Tohoku earthquake, constrained with teleseismic body and surface waves. *Earth, Planets and Space* 63: 559–564.

Shibazaki B, Yoshida Y, Nakamura M, Nakamura M, and Katao H (2002) Rupture nucleations in the 1995 Hyogo-ken Nanbu earthquake and its large aftershocks. *Geophysical Journal International* 149: 572–588.

Simons M, Minson SE, Sladen A, et al. (2011) The 2011 Magnitude 9.0 Tohoku-Oki earthquake: Mosaicking the megathrust from seconds to centuries. *Science* 332: 1421–1425.

Somerville P, Irikura K, Graves R, et al. (1999) Characterizing crustal earthquake slip models for the prediction of strong ground motion. *Seismological Research Letters* 70: 59–80.

Song S, Beroza GC, and Segall P (2005) A Unified Source Model for the 1906 San Francisco earthquake. *Bulletin of the Seismological Society of America* 98: 823–831.

Spudich PKP (1992) On the inference of absolute stress levels from seismic radiation. *Tectonophysics* 211: 99–106.

Spudich P and Frazer LN (1984) Use of ray theory to calculate high-frequency radiation from earthquake sources having spatially variable rupture velocity and stress drop. *Bulletin of the Seismological Society of America* 74: 2061–2082.

Spudich P and Guatteri M (2004) The effect of bandwidth limitations on the inference of earthquake slip-weakening distance from seismograms. *Bulletin of the Seismological Society of America* 94: 2028–2036.

Steidl JH, Archuleta RJ, and Hartzell SH (1991) Rupture history of the 1989 Loma Prieta, California, earthquake. *Bulletin of the Seismological Society of America* 81: 1573–1602.

Suzuki W, Aoi S, Sekiguchi H, and Kunugi T (2012) Source rupture process of the 2011 Tohoku-Oki earthquake derived from the strong-motion records. In: *Proceedings of the 15th Conference on Earthquake Engineering,* p. 1650.

Takeo M (1985) Near-field synthetic seismograms taking into account the effects of anelasticity: The effects of anelastic attenuation on seismograms caused by a sedimentary layer. *Meteorology Geophysics* 36: 245–257 (in Japanese with English abstract).

Takeo M (1987) An inversion method to analyze the rupture processes of earthquakes using near-field seismograms. *Bulletin of the Seismological Society of America* 77: 490–513.

Tanioka Y, Ruff L, and Satake K (1996) The Sanriku-oki, Japan, earthquake of December 28, 1994 (Mw 7.7): Rupture of a different asperity from a previous earthquake. *Journal of Geophysical Research* 23: 1465–1468.

Tarantola A (2005) *Inverse Problem Theory and Methods for Model Parameter Estimation*. Philadelphia, PA: Society of Industrial and Applied Mathematics.

Trifunac MD (1974) A three-dimensional dislocation model for the San Fernando, California, earthquake of February 9, 1971. *Bulletin of the Seismological Society of America* 64: 149–172.

Tsuboi S, Komatitsch D, Ji C, and Tromp J (2003) Broadband modeling of the 2002 Denali fault earthquake on the Earth Simulator. *Physics of the Earth and Planetary Interiors* 139: 305–312.

Uchide T and Ide S (2007) Development of multiscale slip inversion method and its application to the 2004 mid-Niigata Prefecture earthquake. *Journal of Geophysical Research* 112: B06313. http://dx.doi.org/10.1029/2006JB004528.

Uchide T and Ide S (2010) Scaling of earthquake rupture growth in the Parkfield area: Self-similar growth and suppression by the finite seismogenic layer. *Journal of Geophysical Research* 115: B11302. http://dx.doi.org/10.1029/2009JB007122.

Umeda Y (1990) High-amplitude seismic waves radiated from the bright spot of an earthquake. *Tectonophysics* 175: 81–92.

Wald DJ (1996) Slip history of the 1995 Kobe, Japan, earthquake determined from strong motion, teleseismic, and geodetic data. *Journal of the Physics of the Earth* 44: 489–503.

Wald DJ and Graves RW (2001) Resolution analysis of finite fault source inversion using one- and three-dimensional Green's functions. Part 2: Combining seismic and geodetic data. *Journal of Geophysical Research* 106: 8767–8788.

Wald DJ and Heaton TH (1994) Spatial and temporal distribution of slip for the 1992 Landers, California earthquake. *Bulletin of the Seismological Society of America* 84: 668–691.

Wald DJ, Heaton TH, and Hudnut KW (1996) The slip history of the 1994 Northridge, California, earthquake determined from strong-motion, teleseismic, GPS, and leveling data. *Bulletin of the Seismological Society of America* 86: S49–S70.

Wald DJ, Helmberger DV, and Heaton TH (1991) Rupture model of the 1989 Loma Prieta earthquake from the inversion of strong-motion and broadband teleseismic data. *Bulletin of the Seismological Society of America* 81: 1540–1572.

Wald DJ, Kanamori H, Helmberger DV, and Heaton TH (1993) Source study of the 1906 San Francisco earthquake. *Bulletin of the Seismological Society of America* 83: 981–1019.

Wald D and Somerville PG (1995) Variable-slip rupture model of the Great 1923 Kanto, Japan, earthquake: Geodetic and body-waveform analysis. *Bulletin of the Seismological Society of America* 85: 159–177.

Wang D and Mori J (2011) Rupture process of the 2011 off the Pacific coast of Tohoku Earthquake (Mw 9.0) as imaged with back-projection of teleseismic P-waves. *Earth, Planets and Space* 63: 603–607.

Ward SN and Barrientos SE (1986) An inversion for slip distribution and fault shape from geodetic observations of the 1983, Borah Peak, Idaho, earthquake. *Journal of Geophysical Research* 91: 4909–4919.

Wessel P and Smith WHF (1991) Free software helps map and display data. *Earth Observing System Transactions, American Geophysical Union* 72: 441.

Wiemer S and Katsumata K (1999) Spatial variability of seismicity parameters in aftershock zones. *Journal of Geophysical Research* 104: 13135–13151.

Wiens DA (1987) Effects of near source bathymetry on teleseismic P waveforms. *Geophysical Research Letters* 14: 761–764.

Wiens DA (1989) Bathymetric effects on body waveforms from shallow subduction zone earthquakes and application to seismic processes in the Kurile trench. *Journal of Geophysical Research* 94: 2955–2972.

Wu C and Takeo M (2004) An intermediate deep earthquake rupturing on a dip-bending fault: Waveform analysis of the 2003 Miyagi-ken Oki earthquake. *Geophysical Research Letters* 31. http://dx.doi.org/10.1029/2004GL021228.

Wu C, Takeo M, and Ide S (2001) Source process of the Chi-Chi earthquake: A joint inversion of strong motion data and Global Positioning System data with a multifault model. *Bulletin of the Seismological Society of America* 91: 1128–1143.

Yabuki T and Matsu'ura M (1992) Geodetic data inversion using a Bayesian information criterion for spatial distribution of fault slip. *Geophysical Journal International* 109: 363–375.

Yagi Y and Fukahata Y (2008) Importance of covariance components in inversion analyses of densely sampled observed data: An application to waveform data inversion for seismic source processes. *Geophysical Journal International* 175: 215–221.

Yagi Y and Kikuchi M (2000) Source rupture process of Kocaeli, Turkey, earthquake of August 17, 1999, obtained by joint inversion of near-field data and teleseismic data. *Geophysical Research Letters* 27: 1969–1972.

Yagi Y, Kikuchi M, and Sagiya T (2001) Co-seismic slip, postseismic slip, and aftershocks associated with two large earthquakes in 1996 in Hyuga-nada, Japan. *Earth, Planets and Space* 53: 793–803.

Yagi Y, Nakao A, and Kasahara A (2012) Smooth and rapid slip near the Japan Trench during the 2011 Tohoku-oki earthquake revealed by a hybrid back-projection method. *Earth and Planetary Science Letters* 355–356: 94–101.

Yamada T, Mori JJ, Ide S, et al. (2005) Radiation efficiency and apparent stress of small earthquakes in a South African gold mine. *Journal of Geophysical Research* 110. http://dx.doi.org/10.1029/2004JB003221.

Yao H, Gerstoft P, Shearer PM, and Mecklenbräuker C (2011) Compressive sensing of the Tohoku-Oki Mw 9.0 earthquake: Frequency-dependent rupture modes. *Geophysical Research Letters* 38: L20310.

Yao ZX and Harkrider DG (1983) A generalized reflection-transmission coefficient matrix and discrete wavenumber method for synthetic seismograms. *Bulletin of the Seismological Society of America* 73: 1685–1699.

Yao H, Shearer PM, and Gerstoft P (2012) Subevent location and rupture imaging using iterative backprojection for the 2011 Tohoku Mw 9.0 earthquake. *Geophysical Journal International* 190: 1152–1168.

Yokota Y (2013) *Seismo-Geodesy to Infer the Physical Process of the 2011 Tohoku Earthquake*. PhD Thesis, The University of Tokyo.

Yokota Y, Koketsu K, Hikima K, and Miyazaki S (2009) Ability of 1-Hz GPS data to infer the source process of a medium-sized earthquake: The case of the 2008 Iwate-Miyagi Nairiku, Japan, earthquake. *Geophysical Research Letters* 36: L12301. http://dx.doi.org/10.1029/2009GL037799.

Yoshida S (1986) A method of waveform inversion for earthquake rupture process. *Journal of the Physics of the Earth* 34: 235–255.

Yoshida S (1992) Waveform inversion for rupture process using a non-flat seafloor model: Application to 1986 Andreanof Islands and 1985 Chile earthquakes. *Tectonophysics* 211: 45–59.

Yoshida S (1995) Waveform inversion methods for the earthquake source. *Journal of the Physics of the Earth* 43: 183–209.

Yoshida S and Koketsu K (1990) Simultaneous inversion of waveform and geodetic data for the rupture process of the 1984 Naganoken-Seibu, Japan, earthquake. *Geophysical Journal International* 103: 355–362.

Yoshida S, Koketsu K, Shibazaki B, et al. (1996) Joint inversion of near- and far-field waveforms and geodetic data for the rupture process of the 1995 Kobe Earthquake. *Journal of the Physics of the Earth* 44: 437–454.

Yoshida Y, Ueno H, Muto D, and Aoki S (2011) Source process of the 2011 off the Pacific coast of Tohoku earthquake with the combination of teleseismic and strong motion data. *Earth, Planets and Space* 63: 565–569.

Yu S-B, Kuo L-C, Hsu Y-J, et al. (2001) Preseismic deformation and coseismic displacements associated with the 1999 Chi-Chi, Taiwan, earthquake. *Bulletin of the Seismological Society of America* 91: 995–1012.

Yue H and Lay T (2011) Inversion of high-rate (1 sps) GPS data for rupture process of the 11 March 2011 Tohoku earthquake (Mw 9.1). *Geophysical Research Letters* 38: L00G09. http://dx.doi.org/10.1029/2011GL048700.

Zeng Y, Aki K, and Teng TL (1993) Mapping of the high-frequency source radiation for the Loma Prieta earthquake, California. *Journal of Geophysical Research* 98: 11981–11993.

Zeng Y and Anderson JG (1996) A composite source model of the 1994 Northridge earthquake using genetic algorithms. *Bulletin of the Seismological Society of America* 86: S71–S83.

Zeng Y and Chen C-H (2001) Fault rupture process of the 20 September 1999 Chi-Chi, Taiwan, earthquake. *Bulletin of the Seismological Society of America* 91: 1088–1098.

This page intentionally left blank

4.10 Fault Interaction, Earthquake Stress Changes, and the Evolution of Seismicity

GCP King and MH Devès, Institut de Physique du Globe de Paris, Sorbonne Paris Cit.é, Univ. Paris Diderot, Paris, France

© 2015 Elsevier B.V. All rights reserved.

This chapter is a revision of the previous edition chapter by G C P King, Volume 4, pp. 225–255, © 2007, Elsevier B.V.

4.10.1	**Introduction**	243
4.10.2	**Stress Interactions Between Faults in a Homogeneous Elastic Half-Space**	244
4.10.2.1	Coulomb Failure	244
4.10.2.2	2D Case: Coulomb Stress on a Plane of Specified Orientation	245
4.10.2.3	2D Case: Change of Coulomb Stress on Optimally Orientated Faults	246
4.10.2.4	3D Case: Strike-Slip and Dip-Slip Conditions	248
4.10.2.5	Sensitivity to the Main-Shock Focal Mechanism	249
4.10.3	**Examples of Coulomb Interactions**	249
4.10.3.1	Coulomb Stress Changes and Aftershocks	249
4.10.3.2	Stress Changes Associated with the Landers Earthquake	251
4.10.3.3	Coulomb Stress Changes Preceding the Landers Rupture	251
4.10.3.4	Stress Changes Following the Landers Rupture but before the Big Bear Earthquake	251
4.10.3.5	Stress Changes Caused by the Landers, Big Bear, and Joshua Tree Ruptures	252
4.10.3.6	Interactions Between Large Earthquakes: Western Turkey and the Aegean	252
4.10.3.7	Interactions Between Very Large Earthquakes	256
4.10.3.8	Is It Better to Calculate Stress Transfer to Known Faults?	256
4.10.4	**Introducing Time into the Failure Criterion**	257
4.10.4.1	Time-Dependent Rock Failure	257
4.10.4.2	Rate- and State-Dependent Friction Equations	257
4.10.5	**Loading Models and Lithospheric Properties**	258
4.10.5.1	Elastic Half-Space Loading and Stress Relaxation Models	258
4.10.5.2	2D Platelike Models	260
4.10.5.3	Inhomogeneous Lithospheric Models: The Effects of Voids or Fissures	260
4.10.5.4	Interactions Between Volcanic Eruptions and Earthquakes	260
4.10.5.5	Fluid Movement in Nonvolcanic Regions	260
4.10.5.6	The Distribution of Small Events Before Large Earthquakes	262
4.10.5.7	Stress Inhomogeneity Resulting from Displacement Boundary Conditions	262
4.10.6	**Dynamic Triggering**	269
4.10.7	**Conclusions**	269
Acknowledgment		269
References		269

4.10.1 Introduction

This chapter focuses on advances in our understanding of the Earth's lithosphere, over the last few decades – in particular, the way in which studies of earthquakes have added to our knowledge, by shedding light on how earthquakes interact one with another. There is a surprising lack of consensus about how the lithosphere behaves. We know that it is inhomogeneous at scales ranging from microns to tens or hundreds of kilometers, is porous, and is commonly anisotropic – remarks that apply with particular force to the continental lithosphere. Parts of the lithosphere are clearly brittle–elastic and generate earthquakes, but the greater part of the lithosphere is not seismogenic with a mechanical behavior that is poorly resolved. Even the brittle–elastic part behaves strangely when compared with the engineering materials that we normally come across. For example, before half the failure stress is reached in a laboratory sample, it has dilated dramatically. This effect has not been demonstrated to occur at large scales, a disappointment to earthquake prediction attempts in the 1960s and 1970s, but serves to remind us that the behavior of rocks at one scale may be a poor guide to their behavior at another. Commonly, rock sample studies may be only a metaphor for the behavior at larger scales.

How does the nonseismogenic lithosphere behave? Visions vary on this. A persistent view has been that only the upper 10–20 km of the continental lithosphere, the upper crust, is seismogenic (i.e., brittle) and gives way to a lower crust that is viscous with no long-term strength. At greater depth, the Moho is reached and the upper mantle for perhaps 50 km is again strong and brittle.

This view is largely based on the results of laboratory tests conducted on homogeneous samples under very different strain rates, temperature conditions, and boundary conditions from those prevailing in the lithosphere. Nonetheless, the resulting model is mathematically and computationally tractable, and for geophysicists, it has provided a satisfying

explanation for a range of observations. Structural geologists have found the view difficult to accept. While not necessarily claiming that the entire lower crust and mantle are strong, a vision of highly localized shear zones (that only slip above a threshold stress) separated by strong, but deformable, blocks seems more appropriate to describe the structures observed in both active and deeply eroded mountain belts. Unlike the geophysicist's view, this vision is, in general, only mathematically tractable when greatly simplified to a system of rigid blocks separated by narrow deformation zones. These models can be quite complex when describing the mechanics of fold–thrust belts (e.g., Suppe, 1985). The two views have been described respectively as 'dynamic' because they incorporate stresses while the latter are described as kinematic because only geometry (including displacement vectors) is included. An intermediate view includes any time-dependent model based on stresses as dynamic. In this chapter, only phenomena for which momentum terms in their underlying equations cannot be ignored are termed dynamic (seismic waves and dynamic fault rupture).

The two views have been to a large extent supported by the different data sets that the 'geophysicists' and 'structural geologists' customarily study, but both sides can find some discomfort in their simplifications. For example, recent GPS results, best interpreted in terms of blocks, are difficult to reconcile with earlier views that the geodetic data supported distributed deformation. Structural geologists can be perturbed by evidence for massive ductile deformation in the lower crust and upper mantle. Not all geoscientists view the same data; geochemists and hydrologists observed evidence for massive migration of water in the lithosphere, and volcanologists have unequivocal evidence for the existence of massive regions of fluid rock, neither of which figure in the large-scale models outlined.

It was into this garden of confusion that the concepts of Coulomb stress interaction between earthquakes appeared in the 1980s. This met with astonishing success, while making the improbable simplification that the lithosphere is an homogeneous elastic half-space. Using this assumption, the stress on a second fault due to an earthquake on a first can be readily calculated. Section 4.10.2 begins by outlining this simple mathematical formulation for Coulomb stress interaction. Section 4.10.3 then provides examples of the method applied to both aftershock sequences and interactions between large events. Only aftershock sequences for strike-slip faults are illustrated since they offer the most straightforward examples. Section 4.10.4 starts to consider modifications of the simple early theory by the introduction of time-dependent failure through the concept of rate and state friction.

Readers who are already familiar with the basic concepts outlined in Sections 4.10.2–4.10.4 may wish to pass directly to later sections that consider unresolved problems and future perspectives in stress interaction studies. Section 4.10.5 considers the effects of long-term loading and lithospheric inhomogeneities together with the effect of fluids including the interactions between faults and volcanoes. Section 4.10.5 also describes why, in a medium subject to finite strain, the assumptions of infinitesimal strain or stress can never provide a complete description of fault interactions. Finally, the section discusses the accelerated seismic activity that appears over a wide region prior to major earthquakes. Section 4.10.6 touches on the question of dynamic triggering covered fully in Chapter 4.11, and Section 4.10.7 presents concluding remarks.

There are at least 250 papers that can be considered relevant to stress interactions, too many to refer to in any detail. Particular topics are therefore illustrated using a suitable example with no attempt to reference other work on the same topic. For more complete reference lists to earlier papers, the reader is referred to reviews Stein (1999), King and Cocco (2000), Freed (2005), and Steacy et al. (2005).

4.10.2 Stress Interactions Between Faults in a Homogeneous Elastic Half-Space

Faults are widely assumed to be well approximated as very narrow planes across which some components of displacement are discontinuous, while stress components across the plane are continuous. The system is driven by stress boundary conditions that determine the nature of the faulting: strike-slip, normal, or reverse. In turn, the faults exhibit a frictional behavior that determines the stress level at which they will slip to generate an earthquake. As discussed later, each of these assumptions can be challenged, but they allow a simple theoretical and computational framework to be developed on the basis of dislocation theory where static displacements, strains, and stresses can be computed by solving the elastostatic equation for a dislocation on a extended fault in an elastic, isotropic, and homogeneous medium. This solution yields the Volterra equation (Volterra, 1907),

$$u_m(x_i) = \frac{1}{F}\iint_{\Sigma} \Delta u_k(\xi_i) T_{kl}^m(\xi_i, x_i) n_i(\xi_i) \mathrm{d}\Sigma(\xi_i)$$

where F is the magnitude of a volume point force applied at $\vec{\xi}$ in the direction m. n_k are the components of the unit vector normal to the components of the unit vector normal to the surface element $\mathrm{d}\Sigma$. The static displacement $u_m(x_i)$ is computed as a function of the dislocation $\Delta u_k(\xi_i)$ and the static traction $T_{kl}^m(\xi_i, x_i)$ on the fault plane Σ. The previous equation becomes

$$u_m(x_i) = \frac{1}{F}\iint_{\Sigma} \Delta u_k \left[\lambda \delta_{kl} \frac{\partial U_m^n}{\partial \xi_n} + \left(\frac{\partial U_m^k}{\partial \xi_l} + \frac{\partial U_m^l}{\partial \xi_k}\right)\right] n_l \mathrm{d}\Sigma$$

where the summation convention applies (e.g., Aki and Richards, 1980); computation of the static displacement requires knowledge of the fault geometry, the slip distribution, and the strain nuclei (Mindlin, 1936) $U_m^k(\xi_i, x_i)$ (i.e., the static Green's functions). Using these solutions and Hooke's law analytic expressions for the static displacement, strain and stress fields caused by a finite rectangular fault within an elastic half-space have been derived (Okada, 1985).

4.10.2.1 Coulomb Failure

A number of criteria can be used to characterize failure in rocks of which the Coulomb failure criterion is widely used (e.g., Scholz, 1998). It requires that both the shear and normal stress on a preexisting or an incipient fault plane satisfy conditions analogous to those of friction on a preexisting surface.

Failure initiates and spreads on the plane when the Coulomb stress C_f sometimes referred to as the Coulomb failure function (CFF) exceeds a specific value

$$C_f = \tau_\beta + \mu(\sigma_\beta + p) \quad [1]$$

where τ_β is the shear stress on the failure plane, σ_β is the normal stress (positive for extension), p is the pore fluid pressure, and μ the coefficient of friction. The value of τ_β must always be positive in this expression. However, the processes of resolving shear stress onto an assigned plane may give positive or negative values. The difference in sign of τ_β indicates whether the potential for slip on the plane is right or left lateral.

If the failure plane is orientated at β to the σ_1 axis (see Figure 1), we can express, under plane-stress conditions, the normal stress applied to the plane in terms of the principal stresses:

$$\sigma_\beta = \frac{1}{2}(\sigma_1 + \sigma_3) - \frac{1}{2}(\sigma_1 - \sigma_3)\cos 2\beta \quad [2]$$

Two expressions are required for shear stress, the one giving positive values being chosen. One is for left-lateral and the other for right-lateral shear stress:

$$\begin{aligned} \tau_\beta^L &= \frac{1}{2}(\sigma_1 - \sigma_3)\sin 2\beta \\ \tau_\beta^R &= -\frac{1}{2}(\sigma_1 - \sigma_3)\sin 2\beta \end{aligned} \quad [3]$$

where σ_1 is the greatest principal stress and σ_3 is the least principal stress. For left- and right-lateral Coulomb stresses, eqn [1] then becomes:

$$C_f = \frac{1}{2}(\sigma_1 - \sigma_3)(\sin 2\beta - \mu\cos 2\beta) - \frac{1}{2}\mu(\sigma_1 + \sigma_3) + \mu p \quad [4]$$

where the sign of the shear stress is taken to indicate left- or right-lateral motion. By differentiating the last eqn [4] as a function of β, one finds that the maximum Coulomb stress C_f^{max} occurs for two angles when

$$\tan 2\beta = -\mu \quad [5]$$

Pore fluid pressure modifies the effective normal stress across the failure plane as shown in eqn [1]. When rock stress is changed more rapidly than fluid pressure can change through flow, pore fluid pressure, p, can be related to confining stress in the rock by Skempton's coefficient $B(p = -B\sigma_{kk}/3)$, where B varies between 0 and 1. Equation [1] and subsequent expressions (such as eqn [4]) can therefore be rewritten on the assumption that σ_β represents the confining stress as well as the normal stress on the plane ($p = -B\sigma_\beta$)

$$C_f = \tau_\beta + \mu'\sigma_\beta \quad [6]$$

where the effective coefficient of friction is defined by $\mu' = \mu(1 - B)$. Although useful, it must be remembered that the expression is an approximation since the normal stress across a plane needs not be so simply related to confining stress (e.g., Cocco and Rice, 2002). Further problems associated with fluid flow are discussed later.

The failure condition is inherently two-dimensional (2D) with the intermediate stress σ_2 playing no part. Thus, the underlying physics can be illustrated in 2D. To generalize the mathematics to 3D, it is only necessary to determine the orientation of the plane of greatest and least principal stresses in the appropriate coordinate system and to apply the failure conditions in that plane.

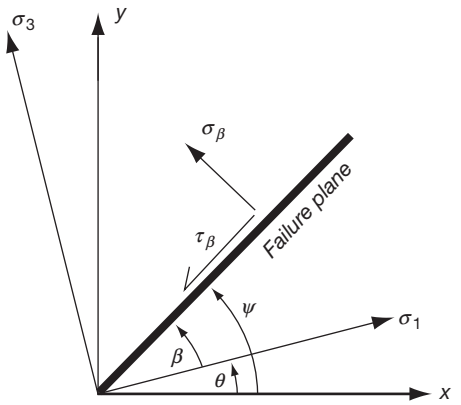


Figure 1 The axis system used for calculations of Coulomb stresses. Extension and left-lateral motion on the failure plane are assumed to be positive. The sign of τ_β is reversed for calculations of right-lateral Coulomb failure on specified planes.

4.10.2.2 2D Case: Coulomb Stress on a Plane of Specified Orientation

In a system where the x- and y-axes and fault displacements are horizontal and fault planes are vertical (containing the z direction), stress on a plane at an angle ψ from the x-axis (see Figure 2) resulting from a general stress field of any origin is given by:

$$\begin{aligned} \sigma_{11} &= \sigma_{xx}\cos^2\psi + 2\sigma_{xy}\sin\psi\cos\psi + \sigma_{yy}\sin^2\psi \\ \sigma_{33} &= \sigma_{xx}\sin^2\psi - 2\sigma_{xy}\sin\psi\cos\psi + \sigma_{yy}\cos^2\psi \\ \tau_{13} &= \frac{1}{2}(\sigma_{yy} - \sigma_{xx})\sin 2\psi + \tau_{xy}\cos 2\psi \end{aligned} \quad [7]$$

where these relations hold for a left-lateral mechanism ($\tau_{13} = \tau_{13}^L$) and similar expressions can be derived for a right-lateral mechanism (as stated in eqn [3]). The Coulomb stress for left-lateral C_f^L and right-lateral C_f^R motion on planes orientated at ψ with respect to the x-axis can now be written in the following way:

$$C_f^L = \tau_{13}^L + \mu'\sigma_{33} \quad [8a]$$

$$C_f^R = \tau_{13}^R + \mu'\sigma_{33} \quad [8b]$$

Equation [8b] is illustrated in Figure 2(a) using a dislocation (earthquake fault) source. An elliptical slip distribution is imposed on the (master) fault in a uniform, stress-free, elastic medium. The contributions of the shear and normal components to the failure condition and the resulting Coulomb stresses for infinitesimal faults parallel to the master fault (commonly referred to as target faults) are shown in separate panels. Such a calculation represents the Coulomb stress on these planes resulting from slip on the master fault (i.e., the figure shows a map of stress changes) and is independent of any knowledge of the prevailing regional stresses or any pre-existing stress fields from other events. The signs in the calculation are chosen such that a positive Coulomb stress indicates a tendency for slip in the same right-lateral sense as the fault of interest. Negative Coulomb stresses indicate a reduction of this tendency. It is important to appreciate that because τ_{13} changes sign between eqns [8a] and [8b], a negative Coulomb stress for right-lateral fault motion is not the same as a tendency for left-lateral slip.

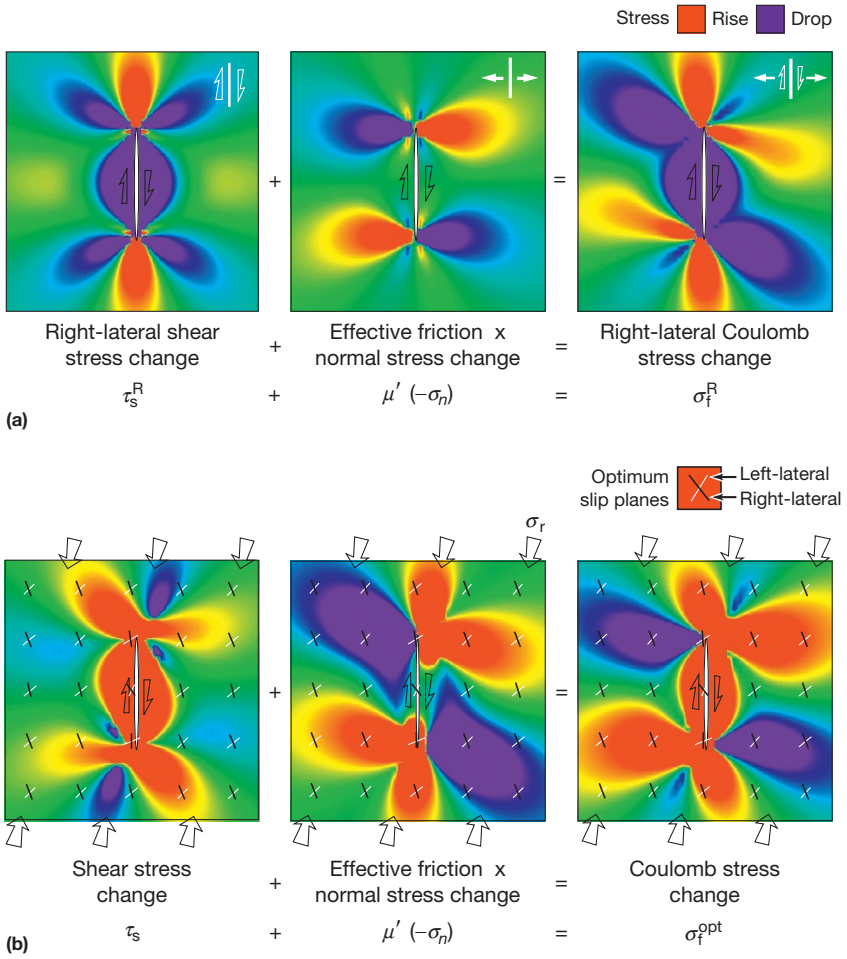


Figure 2 Map view of Coulomb stress changes caused by a right lateral strike-slip fault in plane strain. (a) Coseismic shear, normal, and Coulomb stress changes for right-lateral fault parallel to the master fault. (b) Shear, normal, and Coulomb stress changes on optimally oriented faults (black and white segments) with respect to a regional stress characterized by a N7° E (open arrows) compression of 100 bar. The slip distribution on the fault plane is uniform, but it tapers at the fault ends. The effective friction coefficient μ' is taken to be 0.4. Redrawn after King GCP, et al. (1994b). Static stress changes and the triggering of earthquakes. *Bulletin of the Seismological Society of America* 84: 935–953.

The distribution of increases and decreases of Coulomb stress shows features common to all subsequent figures. Lobes of increased shear stress appear at the fault ends, corresponding to the stress concentrations that tend to extend the fault. Off-fault lobes also appear on either side of the fault, separated from the fault by a region where the Coulomb stresses have not been increased. If the master fault was infinitesimal in length and thus behaving as a point source, all four lobes would be equal in amplitude to the fault-end lobes at all distances. The normal stress field is similar to the more familiar dilatational field with maxima and minima distributed antisymmetrically across the fault, but we consider only the component of tension normal to the fault. The influence of the normal stress on the Coulomb stress distribution is to reduce the symmetry of the final distribution and to increase the tendency for off-fault failure.

4.10.2.3 2D Case: Change of Coulomb Stress on Optimally Orientated Faults

The calculation in the previous section is general and need not be due to an earthquake. If a stress or stress change is known, then the Coulomb stress or stress change can be calculated on any plane. The more general case where a change of stress, such as that due to an earthquake, is imposed on a preexisting stress field can be calculated. The failure planes are not specified a priori, but are computed accounting for the interaction between the two fields. Optimal planes are regarded as those simultaneously favored by tectonic loading and the stress induced by the motion on the master fault. Here, we define such planes in terms of uniformly varying stress fields. The optimum directions are determined by the stress change due to that earthquake $\sigma_{ij}^{q}(=\Delta\sigma_{ij})$ and by existing regional stresses σ_{ij}^{r} that gives a total stress σ_{ij}^{t}

$$\sigma_{ij}^{t} = \sigma_{ij}^{r} + \sigma_{ij}^{q} \qquad [9]$$

The orientation of the principal axes resulting from the total stress is therefore derived using

$$\theta = \frac{1}{2}\tan^{-1}\left(\frac{2\sigma_{xy}^{t}}{\sigma_{xx}^{t} - \sigma_{yy}^{t}}\right) \qquad [10]$$

where θ is the orientation of one principal axis to the x-axis as shown in Figure 1 and the other is at $\theta \pm 90°$. From these two

directions, the angle of greatest compression θ_1 must be chosen. The optimum failure angle ψ_0 is then given by $\theta_1 \pm \beta$ where one plane is left and the other right lateral with the sign of shear stress being chosen appropriately. Whereas the optimum planes are determined from σ_{ij}^{t}, the normal and shear stress changes on these planes are determined only by the earthquake stress changes σ_{ij}^{q}. Thus, the changes in stress on the optimum planes become

$$\begin{aligned}\sigma_{33} &= \sigma_{xx}^{q}\sin^2\psi_0 - 2\sigma_{xy}^{q}\sin\psi_0\cos\psi_0 + \sigma_{yy}^{q}\cos^2\psi_0 \\ \tau_{13} &= \frac{1}{2}\left(\sigma_{yy}^{q} - \sigma_{xx}^{q}\right)\sin 2\psi_0 + \tau_{xy}^{q}\cos 2\psi_0\end{aligned} \quad [11]$$

and the Coulomb stress changes on the planes are equal and given by

$$C_f^{opt} = \tau_{13} + \mu'\sigma_{33} \quad [12]$$

The two optimum planes correspond to left- and right-lateral shear. The Coulomb stress change is the same on both so that the Coulomb stress is the same whether calculated for left- or right-lateral planes in expression [12]. It is important to emphasize that we calculate the change of Coulomb stress on planes that are optimum after the earthquake with the optimum orientations being calculated from the earthquake stress field plus the regional stress field. The Coulomb stress changes caused by the earthquake stress changes are then resolved onto these planes. Where earthquake stresses are large, they can rotate the principal axes.

The results of a calculation to find optimum orientations and magnitudes of Coulomb stress changes are shown in Figure 2(b). The slip on the master fault is the same as before (Figure 2(a)). A uniform 100-bar compressional stress is introduced with the orientation shown. White lines indicate optimum left-lateral orientations and black lines, right-lateral orientations. The shear and normal stress contributions to the Coulomb stress change are again shown in separate panels.

It can be seen from expression [9] that only the deviatoric part of the regional stress determines the orientation of principal axes and hence the optimum stress orientations. Thus, it is sufficient to apply the regional stress as a simple uniaxial compression or extension. This assumes, however, the intermediate principal stress is vertical and thus a 2D description is complete. In general, this is not true since the relative magnitudes of the principal stresses control the focal mechanisms that need not be strike-slip.

The relative amplitude of the regional stress to the earthquake stress drop ($\Delta\tau|$) might be expected to have an effect. This is explored for the strike-slip case in Figure 3, which shows the Coulomb stress change C_f^{opt} on optimally orientated right-lateral planes in which the regional field is equal to the stress drop $\Delta\tau|$ (*left panel*) and ten times $\Delta\tau|$ (*right panel*). These examples span likely conditions. It is evident that, except close to the master fault, the orientations of the optimal planes and Coulomb stress changes on these planes are little altered. The optimal orientations are essentially fixed by the regional stress except very close to the fault where the stress change caused by slip on the master fault is comparable to the regional stress. Thus, for stress drops similar in amplitude to the regional stress, Coulomb stress changes close to the fault are always positive. (For a regional stress of zero, they are positive everywhere.)

The effects of varying the orientation of regional stresses and changing the coefficient of friction μ' are shown in Figure 4. Possible changes of regional stress orientation are limited since the main fault must move as a result of the

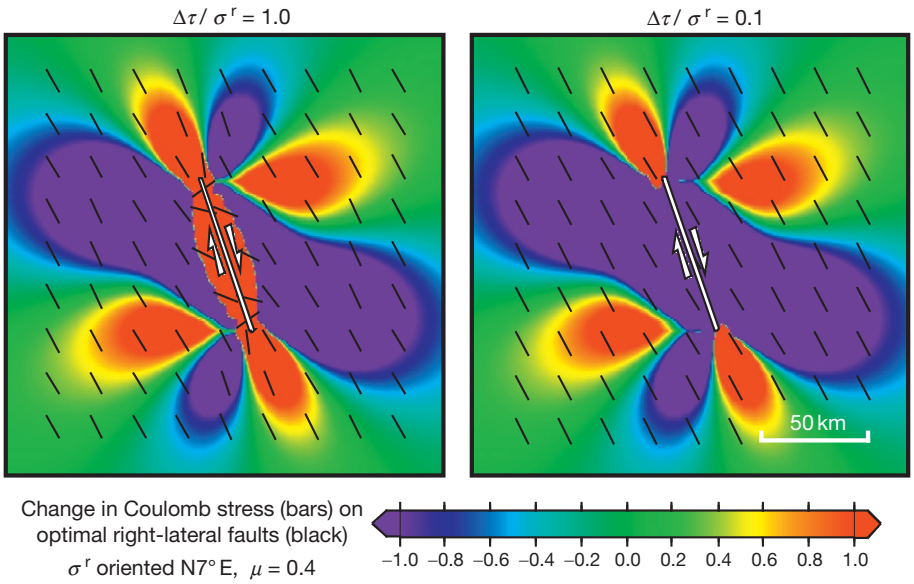

Δτ/ σ r = 1.0
Δτ/ σ r = 0.1
50 km
Change in Coulomb stress (bars) on optimal right-lateral faults (black)
σ r oriented N7° E, μ = 0.4
–1.0 –0.8 –0.6 –0.4 –0.2 0.0 0.2 0.4 0.6 0.8 1.0

Figure 3 Coulomb stress changes for different magnitudes of the regional stress. The left panel shows the Coulomb stress changes for an earthquake stress drop similar in amplitude to the deviatoric regional stress, while the right panel shows the same calculations for a larger (ten times) regional stress amplitude. The former case represents the stress perturbation caused by an earthquake relieving all of the regional stress. In this case, the resulting optimally oriented planes (black segments) for failure rotate from the direction predicted by the regional deviatoric stress near the causative fault, where the induced coseismic stress is larger. Far away from the fault (left panel), where the regional stress amplitude is larger than the induced one, the optimally oriented planes follow the regional stress (see also the right panel). Redrawn after King GCP, et al. (1994b) Static stress changes and the triggering of earthquakes. *Bulletin of the Seismological Society of America* 84: 935–953.

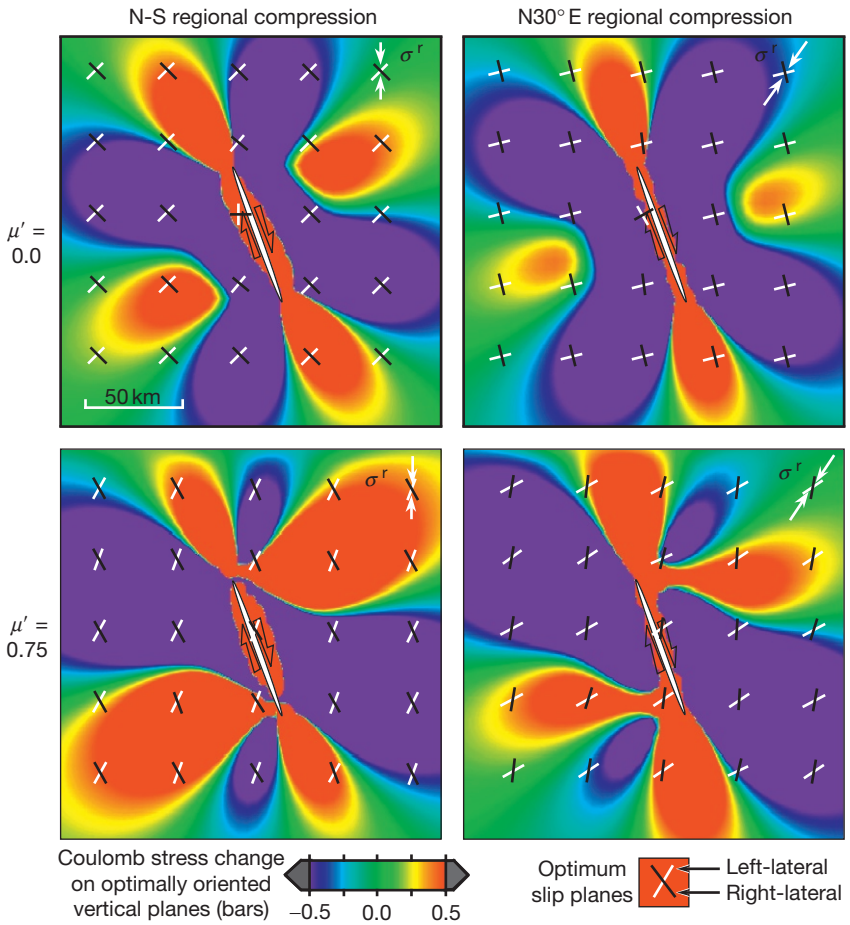


Figure 4 Coulomb stress changes for different orientations of the regional deviatoric stress (left and right panels) and for different values of the effective friction coefficient (upper and bottom panels). Both configurations of the regional stress field refer to horizontal s_1 and s_3 directions. Stress changes are caused by a 70 km long and 12.5 km wide vertical right-lateral fault with 5 m of tapered slip. The regional stress amplitude is 100 bar as in **Figure 3**. The regional stress orientation controls the size of the off-fault to fault-end lobes. The effective friction coefficient controls the internal angle between right- and left-lateral slip planes. Redrawn after King GCP, et al. (1994b) Static stress changes and the triggering of earthquakes. *Bulletin of the Seismological Society of America* 84: 935–953.

regional stress; the 30° range covers the likely range. Similarly, values of friction between 0.0 and 0.75 span the range of plausible values. All of the panels show the same general features, fault-end and off-fault Coulomb stress lobes. Thus, modeling is most sensitive to the regional stress direction, almost insensitive to the regional stress amplitude and modestly sensitive to the coefficient of effective friction. In other words, while the relative magnitude of the regional stress with respect to the earthquake stress drop controls the changes in the orientations of the optimal planes close to the master fault (as shown in **Figure 4**), the orientation of the regional stress is the most important factor controlling the orientations of the optimal planes for failure far away from the master fault.

4.10.2.4 3D Case: Strike-Slip and Dip-Slip Conditions

The application to vertical strike-slip faults is easy because a 2D problem (plane-stress configuration) can be solved with the vertical components of the regional stress tensor neglected. However, the application to dip-slip faults requires the solution of a 3D problem where the ratio between vertical and horizontal components of regional stress tensor must be known. This means that, while in the foregoing discussion, we have assumed that only strike-slip vertical faults are present and thus stress components $\sigma_{zz}, \sigma_{xz}, \sigma_{yz}$ could be neglected, in a more general configuration, these stress components cannot be ignored.

If the regional stress components are known, the total stress can be computed through eqn [9], which is used to find the orientation and magnitude of the principal stresses. We can therefore calculate the orientation of the plane containing σ_1 and σ_3, and hence, the optimum orientations of slip planes where the change of Coulomb failure stress will be found. The calculation is straightforward and an example is shown in **Figure 5**. If we take horizontal stresses similar in amplitude to those employed before, we can examine the change of mechanism as the vertical stress is changed. The horizontal stresses are chosen to be 200 (EW) and 400 bar (NS). For **Figure 5(a)**, the vertical stress is 100 bar; in **Figure 5(b)**, the vertical stress is 300 bar; and in **Figure 5(c)**, the vertical stress is 500 bar. The predominant mechanisms are reverse faulting in **Figure 5(a)**, strike-slip faulting in **Figure 5(b)**, and normal

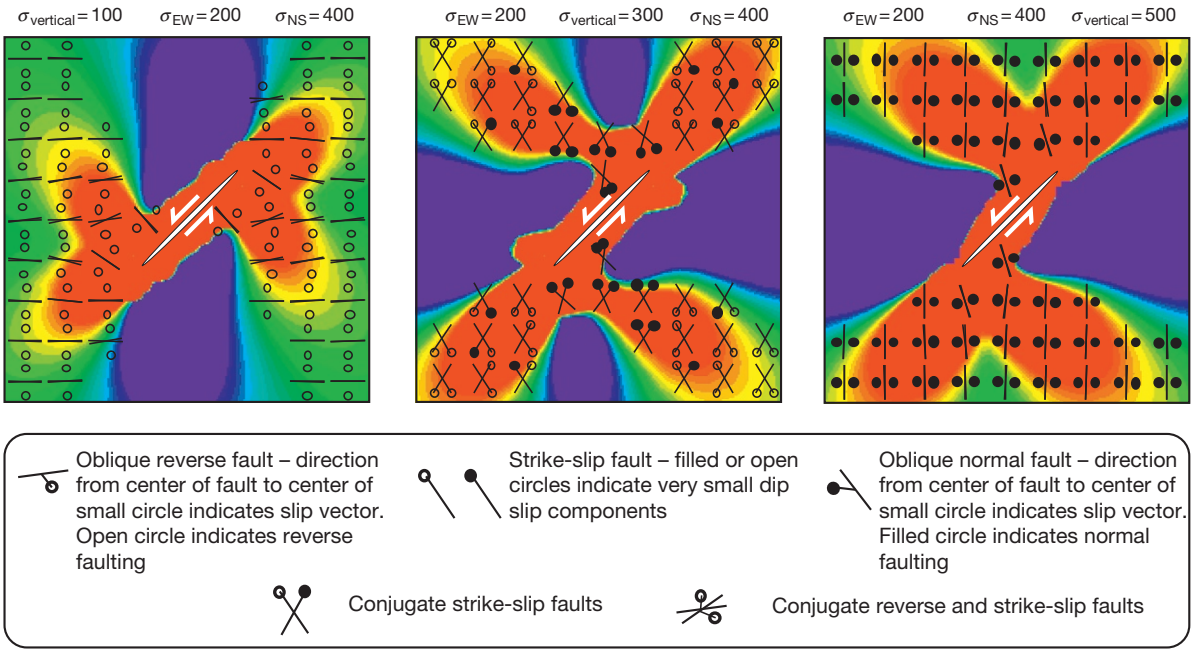

$\sigma_{vertical} = 100$
$\sigma_{EW} = 200$
$\sigma_{NS} = 400$
$\sigma_{EW} = 200$
$\sigma_{vertical} = 300$
$\sigma_{NS} = 400$
$\sigma_{EW} = 200$
$\sigma_{NS} = 400$
$\sigma_{vertical} = 500$
Oblique reverse fault – direction from center of fault to center of small circle indicates slip vector. Open circle indicates reverse faulting
Strike-slip fault – filled or open circles indicate very small dip slip components
Oblique normal fault – direction from center of fault to center of small circle indicates slip vector. Filled circle indicates normal faulting
Conjugate strike-slip faults
Conjugate reverse and strike-slip faults

Figure 5 Coulomb stress changes and optimally slip planes for different orientations of the regional stress field. The vertical stress coincides with the least principal stress (100 bar) in the left panel, with the intermediate principal stress (300 bar) in the middle panel, and with the greatest principal stress (500 bar) in the right panel. Therefore, the predicted faulting mechanisms on optimally oriented planes are strike-slip in the middle panel and inverse or normal in the left and right panels, respectively.

faulting in Figure 5(c). If the vertical stresses are attributed to overburden pressure, then the figures correspond to depths of 300, 900, and 1500 m, respectively. This is clearly incorrect. Events do not occur at such shallow depths nor are systematic changes of focal mechanism with depth observed. It would seem that a confining pressure that increases with depth must be added to the horizontal stress such that the differential stresses remain similar. It is clear, however, that the mechanisms are sensitive to small variations in the stresses chosen; thus, to correctly predict mechanisms from stresses, the stress regime must be known with precision. Direct measurements of stress throughout the seismogenic zone are never available; thus, predicting mechanisms from stress must be indirect. In practice, the most useful stress information comes from some sort of average of the stress orientation required to generate the observed focal mechanisms in the region. As discussed in Section 4.10.3.8, we explain that calculation of stress changes on individual fault planes from determined focal mechanisms might be preferred.

4.10.2.5 Sensitivity to the Main-Shock Focal Mechanism

The last section showed that the Coulomb distribution can be sensitive to incorrect assumptions about regional stress since this can cause target fault focal mechanisms to be incorrectly predicted. While nearby events are clearly sensitive to all of the details of the rupture process, this is not the case some distance from the fault. As shown in Section 4.10.2.3, the orientation of the regional stress is the most important factor controlling the orientations of the optimal planes for failure far away from the master fault. This is particularly important for studying fault interaction between large magnitude earthquakes. Completely incorrect focal mechanisms clearly result in incorrect results. Errors in dip (within an acceptable range) are much less serious. Figure 6(a) shows the distribution for a 45° dipping normal fault, and Figure 6(b) shows the distribution for a vertical dike with the same opening displacement as the horizontal slip vector of the dip-slip fault. Except at distances comparable to the source dimensions, the two distributions are identical. This illustrates that, except close to a dip-slip fault, the fault dip is not important. In general, except in the near field (or at distances of few fault lengths), a double couple point source with the appropriate seismic moment is all that is needed.

While large errors in the strike of the focal planes are serious, small errors only result in (an approximately) commensurate rotation of the Coulomb stress distribution. This is shown in Figure 6(c). It can be remarked at this point that it is much harder to evaluate the fault parameters close to dip-slip faults making aftershock studies of such events more difficult than for strike-slip events. We therefore choose to illustrate aftershock studies for strike-slip events and large event interactions for dip-slip events.

4.10.3 Examples of Coulomb Interactions

4.10.3.1 Coulomb Stress Changes and Aftershocks

The methods outlined earlier can be applied to aftershock distributions. Figure 7 shows the 1979 Homestead Valley aftershock sequence. There are four lobes of Coulomb stress rise and four lobes of Coulomb stress drop. The lobes at the ends of the fault extend into the fault zone, while the off-fault lobes are separated from the fault over most of its length by a

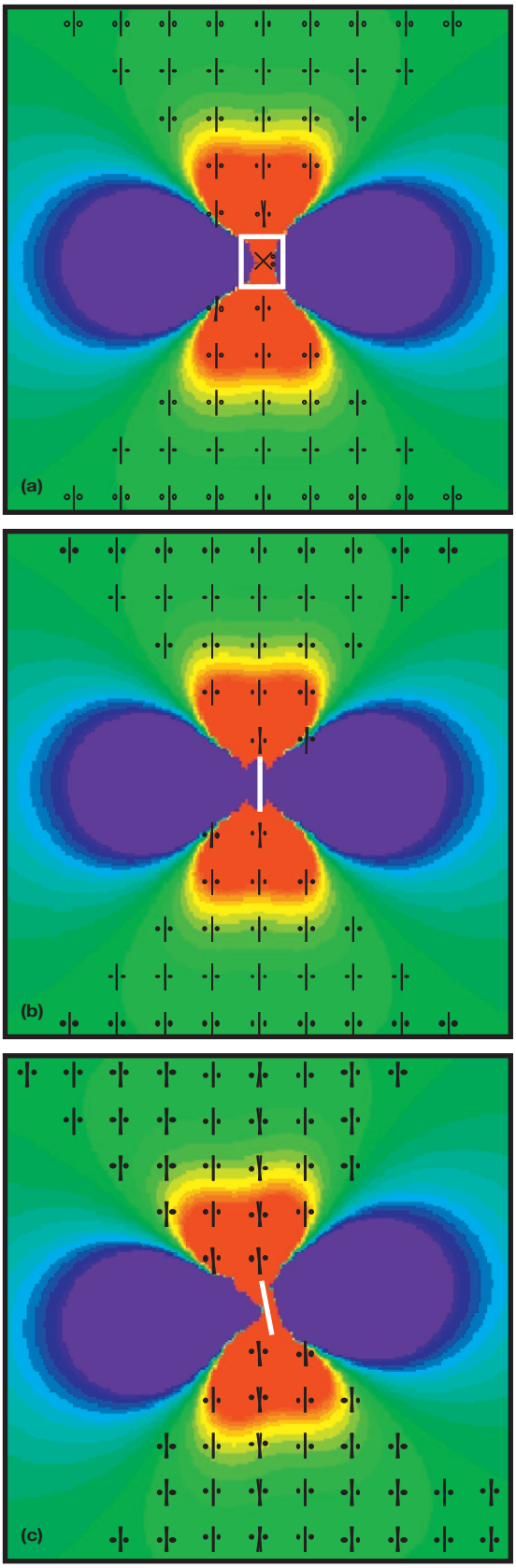


Figure 6 Coulomb stress changes caused by a 45° dipping normal fault (a), a vertical dike with the same opening displacement as the horizontal slip vector of the dip-slip fault (b), and a 45° dipping normal fault with a different strike (c). Except at distances comparable to the source dimensions, the two distributions are identical, suggesting that, except close to a dip-slip fault, the fault dip is less important than fault strike; in this latter case, errors only result in an approximately commensurate rotation of the Coulomb stress distribution. Symbols as in Figure 5.

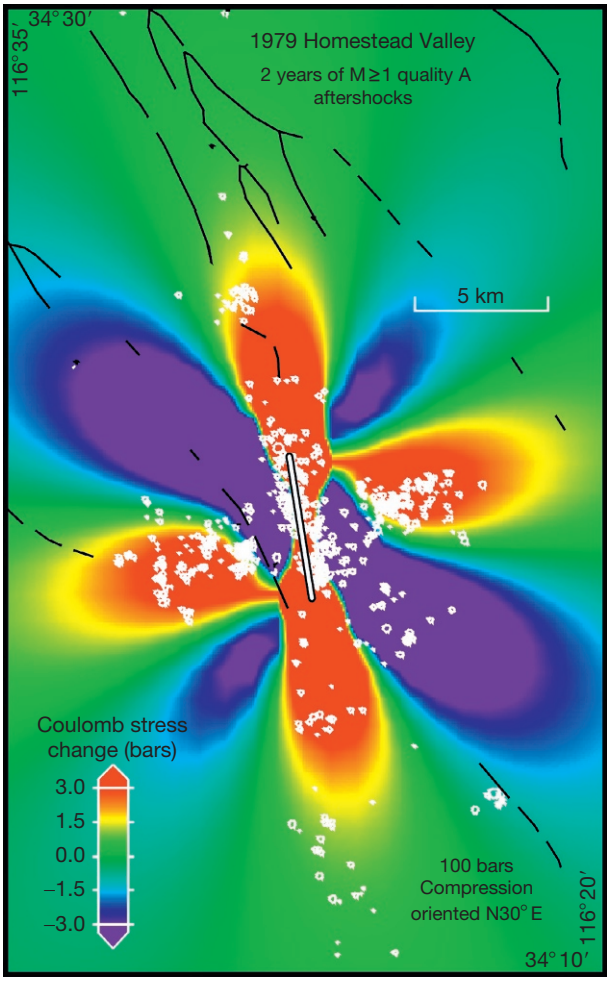


Figure 7 Coulomb stress changes caused by the 15 March 1979 Homestead Valley earthquake sequence (redrawn after King GCP, et al. (1994b) Static stress changes and the triggering of earthquakes. *Bulletin of the Seismological Society of America* 84: 935–953) and aftershock distribution (the same shown in Figure 1). The calculations are carried out in a half-space with the values of Coulomb stress plotted in the figures being calculated at half the depth to which the faults extend. The fault is 5.5 km long and 6 km wide and slip is 0.5 m and it is tapered at the fault ends.

zone where the Coulomb stress is reduced. The distributions of aftershocks are consistent with these patterns. Many events are associated with increases of Coulomb stress of <1 bar, while reductions of the same amount apparently suppress them. Relatively few events fall in the regions of lowered Coulomb stress, and the clusters of off-fault aftershocks are separated from the fault itself by a region of diminished activity. The distributions of Coulomb stresses can be modified as described earlier by adjusting the regional stress direction and

changing μ. However, any improvements in the correlation between stress changes and aftershock occurrence are modest. Consequently, we have chosen to show examples with an average μ' of 0.4. Whatever values we adopt, we find that the best correlations of Coulomb stress change to aftershock distribution are at distances greater than a few kilometers from the fault. Closer to the fault, unknown details of fault geometry and slip distribution influence the stress changes. Correlation between aftershock distribution and Coulomb stress changes on a vertical cross section can also be observed and are discussed by King et al. (1994b).

4.10.3.2 Stress Changes Associated with the Landers Earthquake

Although stress interactions had been previously observed for other events, the clear correlations associated with the Landers earthquake suggested to the scientific community that Coulomb calculations might indeed prove an effective method of relating large events with each other and relating large events to their aftershock sequences. The Landers earthquake and associated events remain the best example for illustrating the techniques and we reproduce here the modeling processes following King et al. (1994b). Since the main events were strike-slip (except for Cedar Mountain), the calculations are essentially 2D and an 'effective' regional stress is easy to establish since only the deviatoric part in the horizontal plane is needed.

4.10.3.3 Coulomb Stress Changes Preceding the Landers Rupture

Figure 8 shows the Coulomb stress changes caused by the four $M > 5$ earthquakes within 50 km of the epicenter that preceded the Landers earthquake. The Coulomb stresses are due to the 1975 $M_L = 5.2$ Galway Lake, 1979 $M_L = 5.2$ Homestead Valley, 1986 $M_L = 6.0$ North Palm Springs, and 1992 $M_L = 6.1$ Joshua Tree earthquakes. These progressively increased Coulomb stresses by about 1 bar at the future Landers epicenter. Together, they also produced a narrow zone of Coulomb stress increase of 0.7–1.0 bar, which the future 70 km long Landers rupture followed for 70% of its length. The Landers fault is also nearly optimally oriented for most of its length. It is noteworthy that the three largest events are roughly equidistant from the future Landers epicenter: the right-lateral Homestead Valley and Joshua Tree events enhanced stress as a result of the lobes beyond the ends of their ruptures, whereas the North Palm Springs event enhanced rupture as a result of an off-fault lobe. Increasing the effective friction μ' from 0.4 (Figure 8) to 0.75 slightly enhances the effects and dropping the friction to zero reduces them.

4.10.3.4 Stress Changes Following the Landers Rupture but before the Big Bear Earthquake

Unlike the earthquake sources modeled so far, which are approximated by an elliptical slip distribution on a single plane, there is more information about the $M = 7.4$ Laders source that following the rupture is modeled with 13 fault

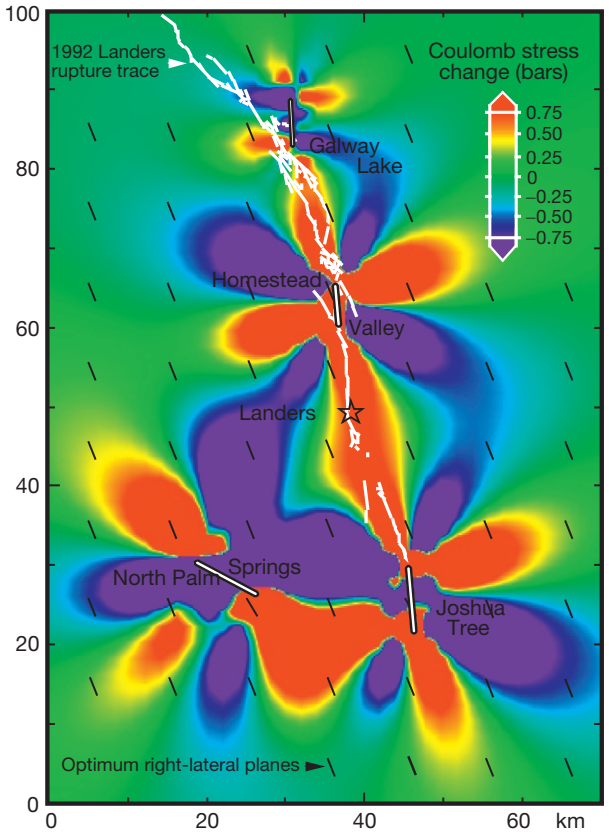


Figure 8 Coulomb stress changes caused by the four earthquakes that preceded the 1992 Landers main shock within 50 km of the epicenter. The source parameters used in this calculation are listed in King et al. (1994b). These events raised the stress at the future Landers epicenter and on most (70%) of the Landers rupture.

segments to produce a slip distribution that is consistent with surface fault mapping, geodetic data, radar data, and the modeling of seismic data.

The stress changes caused by the Landers event are shown in Figure 9. The largest lobe of increased Coulomb stress is centered on the epicenter of the future $M_L = 6.5$ Big Bear event, where stresses were raised 2–3 bar. The Big Bear earthquake was apparently initiated by this stress rise 3 h 26 min after the Landers main shock. The Coulomb stress change at the epicenter is greatest for high effective friction but remains more than 1.5 bar for $\mu' = 0$. There is no surface rupture or Quaternary fault trace associated with the Big Bear earthquake. Judging from its epicenter and focal mechanism, the plane that apparently ruptured was optimally aligned for left-lateral failure, with the rupture apparently propagating northeast and terminating where the Landers stress change became negative.

In addition to calculating the stress changes caused by the Landers rupture, an estimate can be made of the slip on the Big Bear fault needed to relieve the shear stress imposed by the Landers rupture. This is achieved by allowing the Big Bear rupture plane to slip freely as in the stress field caused by the Landers rupture. The potential slip along the Big Bear fault is 60 mm (left lateral), about 5–10% of the estimated slip that

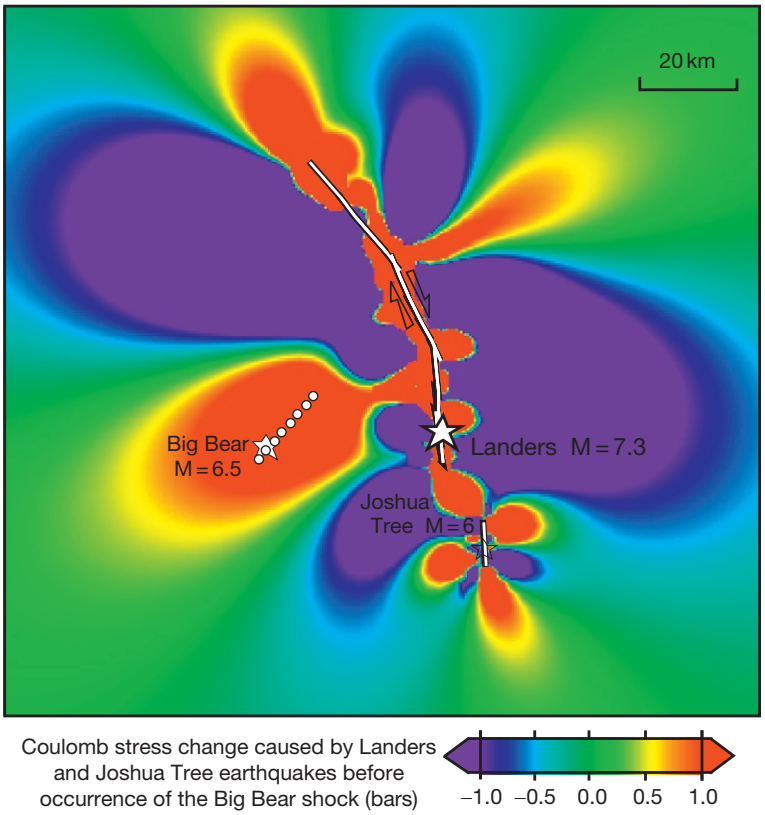


Figure 9 Coulomb stress changes caused by the 1992 Joshua Tree and Landers earthquakes. The location of the Big Bear aftershock and its faulting mechanism are consistent with the induced coseismic stress field. Redrawn after King GCP, et al. (1994b) Static stress changes and the triggering of earthquakes. *Bulletin of the Seismological Society of America* 84: 935–953.

occurred. These calculations suggest that the Big Bear slip needed to relieve the stress imposed by Landers was a significant fraction of the total slip that later occurred. Thus, from consideration of the stress changes and the kinematic response to those changes, it has been considered reasonable to propose that stresses from the Landers event played a major role in triggering the Big Bear shock.

4.10.3.5 Stress Changes Caused by the Landers, Big Bear, and Joshua Tree Ruptures

The Big Bear earthquake was the largest of more than 20 000 aftershocks located after the Landers earthquake, large enough to result in significant stress redistribution at the southwestern part of the Landers rupture zone. Consequently, the distribution of later events cannot be examined without considering its effect. Although smaller, a similar argument can be applied to the Joshua Tree event, whose aftershock sequence was not complete at the time of the Landers rupture. In Figure 10, the combined Coulomb stress changes for the Joshua Tree, Landers, and Big Bear earthquakes are plotted. This distribution is shown together with all well-located $M_L \geq 1$ events that occurred during the following 25 days.

Most $M_L > 1$ aftershocks occur in regions where the failure stress is calculated to have increased by ≥ 0.1 bar, and few events are found where the stress is predicted to have dropped (Figure 10). Even when all seismicity within 5 km of the Landers, Big Bear, and Joshua Tree faults is excluded, more than 75% of the aftershocks occur where the stress is predicted to have risen by >0.3 bar. In contrast, $<25\%$ of the aftershocks occur where the stress dropped by >0.3 bar.

4.10.3.6 Interactions Between Large Earthquakes: Western Turkey and the Aegean

The Landers earthquake sequence suggests that a series of smaller events that preceded the Landers earthquake prepared a region of slightly elevated stress along much of the fault that the main event then followed (Figure 8). It is also clear that the stresses created by the main event largely controlled the aftershock distribution, including the location of the large 'Big Bear' aftershock. A number of studies have shown that large events appear to interact over large areas (see reviews quoted in the introduction for more detailed reference lists). Here, we illustrate the effect taking an example from Turkey and the western Aegean. Since 1912, 29 events ($M_s \geq 6.0$) have occurred in the region shown in Figure 11. The area in the west is of particular interest because the events, unlike the sequence that occurred further east on the North Anatolian Fault events, do not lie along a single fault requiring to result from both on- and off-fault Coulomb stress changes. The events are predominantly strike-slip and normal faulting, but some reverse faulting occurs. The later events have reliable focal mechanisms and seismic moments controlled by waveform modeling. Earlier

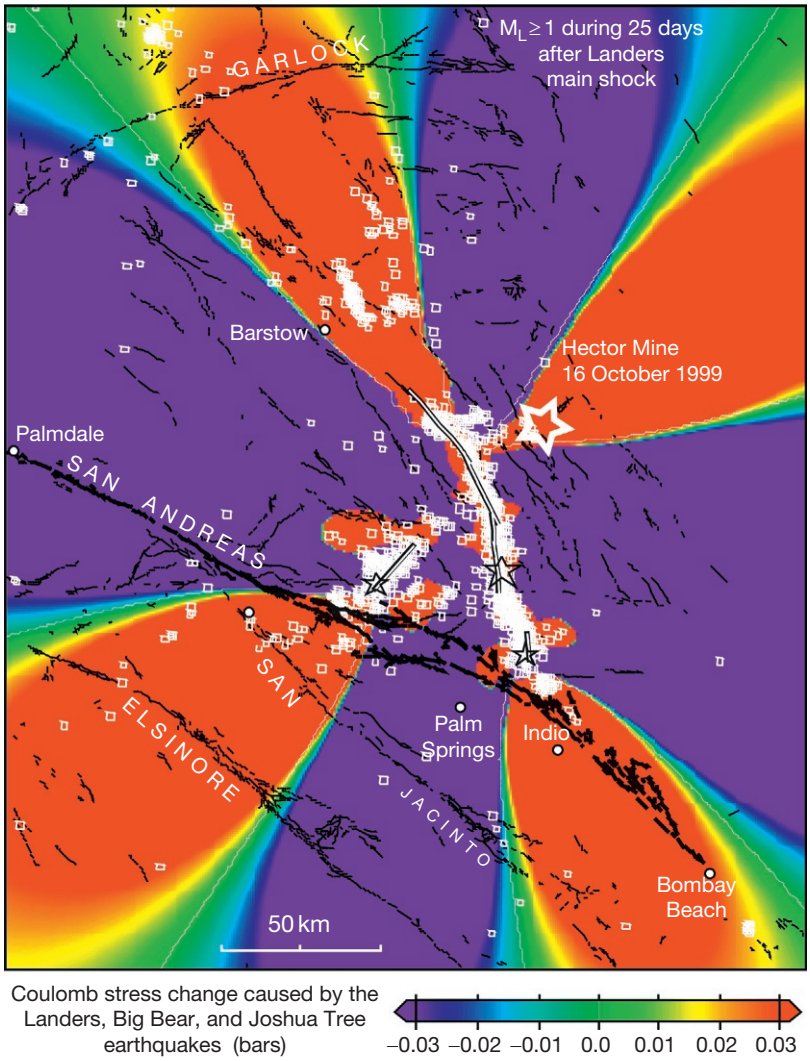
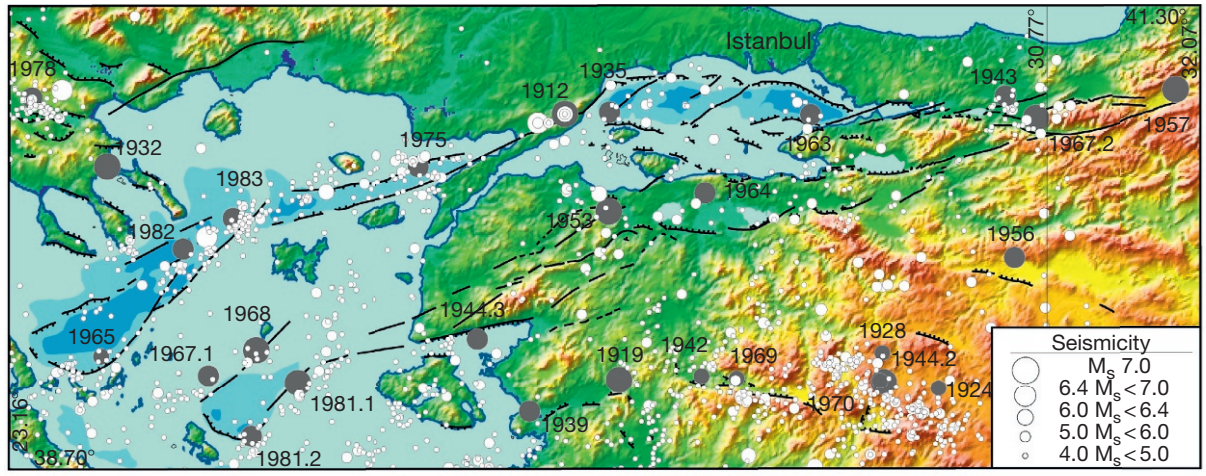
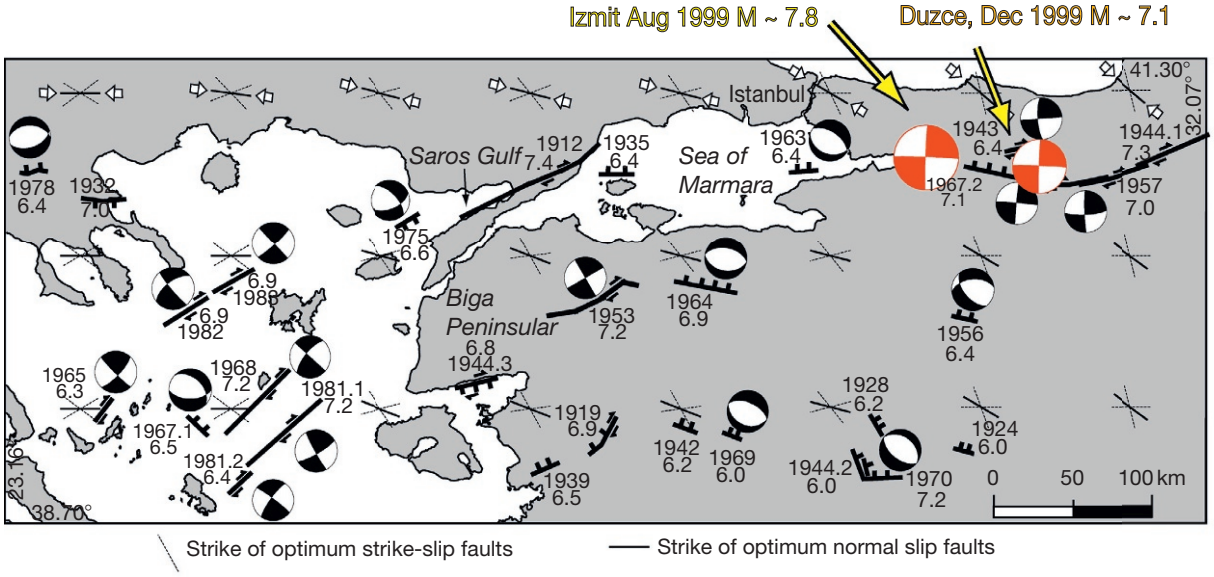


Figure 10 Coulomb stress changes caused by the Joshua Tree, Landers, and Big Bear earthquakes mapped at a depth of 6.25 km (halfway down the fault) and ($M \geq 1$) aftershock distribution during the 25 days following the Landers main shock. Source parameters are those used by King et al. (1994a,b). The location (open star) and the surface rupture (solid black line) of the M 7.1, 1999, Hector Mine earthquake epicenter are shown on the map. The Coulomb stress changes are plotted with a reduced amplitude (color) scale with respect to those of the previous figures. Redrawn after King GCP, et al. (1994b) Static stress changes and the triggering of earthquakes. *Bulletin of the Seismological Society of America* 84: 935–953.

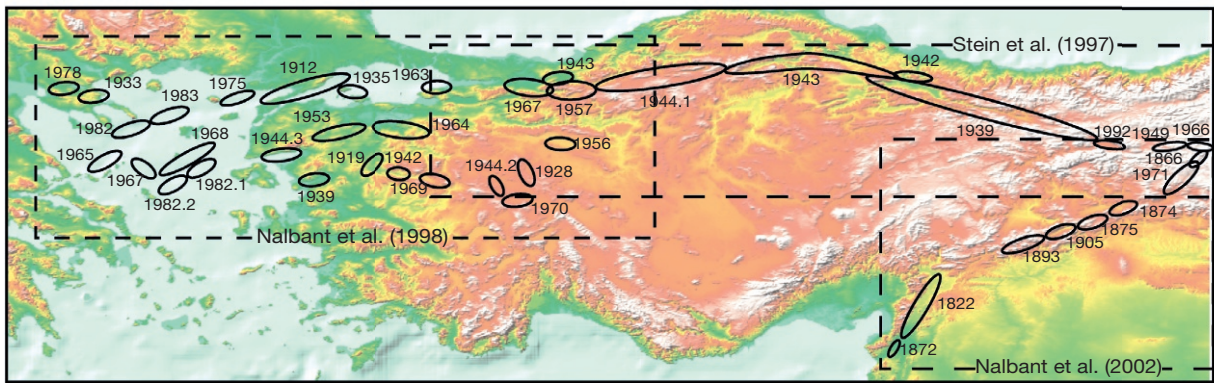
events do not, but a combination of geologic information to identify the faults involved together with damage information closely controls the possible mechanisms. Details of the mechanism information can be found in Nalbant et al. (1998) from which the examples presented here are abstracted. That paper splits the time period into nine intervals such that all interactions can be seen. Here, we only show three example time windows: they are 1912–44 (**Figure 12(a)**), 1912–67 (**Figure 12(b)**), and 1912–96 (**Figure 12(c)**). These are insufficient to demonstrate all the interactions that require the original figures to be examined but allow the reader to gain an impression of the size of the events and the distances between them. The results can be summarized as follows: out of 29 events, 16 (1935, 1939, 1944.1, 1953, 1957, 1964, 1967.2, 1968, 1969, 1970, 1975, 1978, 1981.1, 1981.2, 1982, and 1983) occurred in regions where stress was increased by more than 0.1 bar. If less certain information about earlier events is used and smaller stress increases are presumed to be significant, then 23 of the 29 events appear to have responded to Coulomb stress interactions. Stein et al. (1997) examined interactions between ten events extending to the east of the region shown in **Figure 10(c)**. Of these, all but one (eight out of ten) occurred in regions of substantially enhanced Coulomb stress. For the East Anatolian Fault, Nalbant et al. (2002) (**Figure 10(c)**) had shown that nine out of ten events ($M > 6.7$) occurred in regions of enhanced Coulomb stress. Taken together with the Nalbant et al. (1998) work, 33 out of 49 events since 1912 along the North and East Anatolian fault and the Aegean system show clear Coulomb interactions. If less certain information is included, then this rises to 41 out of 49. Of great importance, not one substantial event occurred in the mauve regions where Coulomb stress was reduced. Such regions have been described as stress shadows (Harris, 1998; Toda et al., 2012) and with some reservations discussed later



(a)

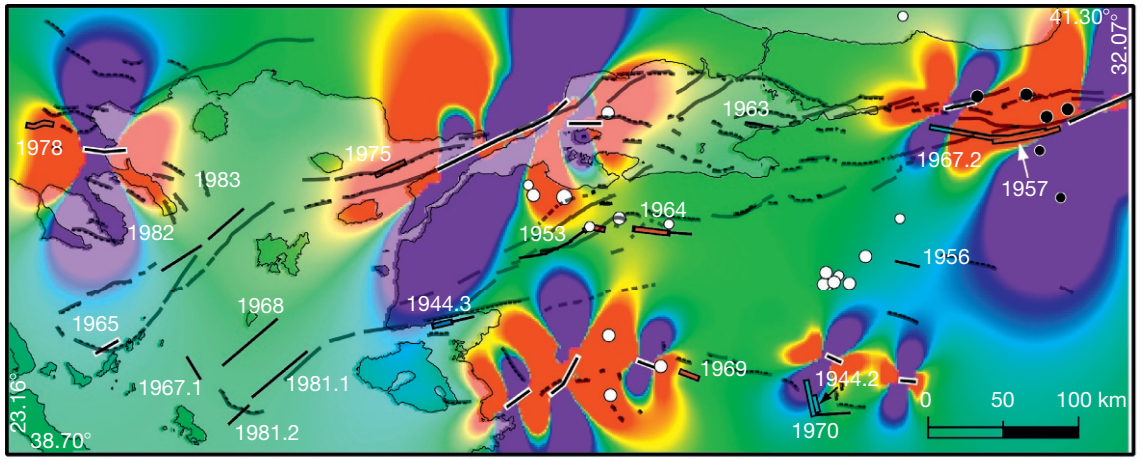


(b)

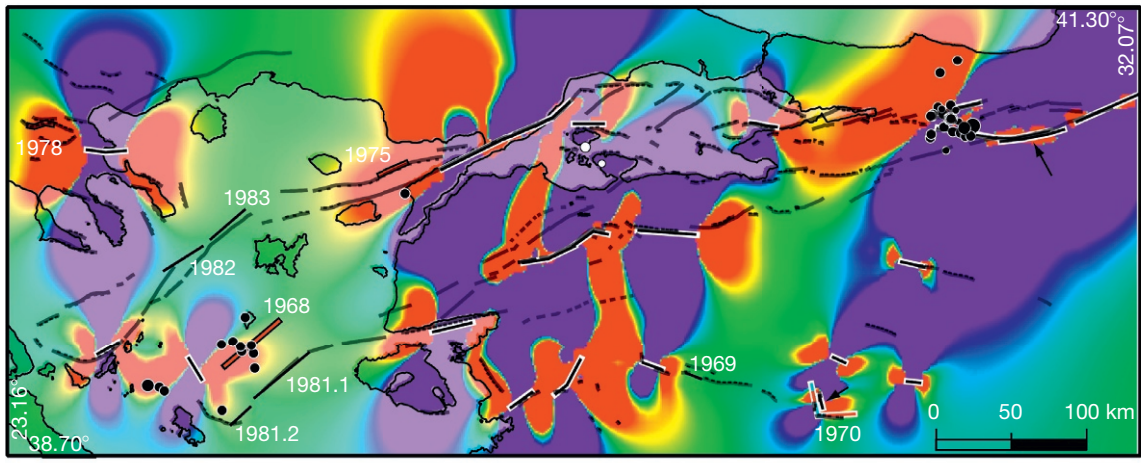


(c)

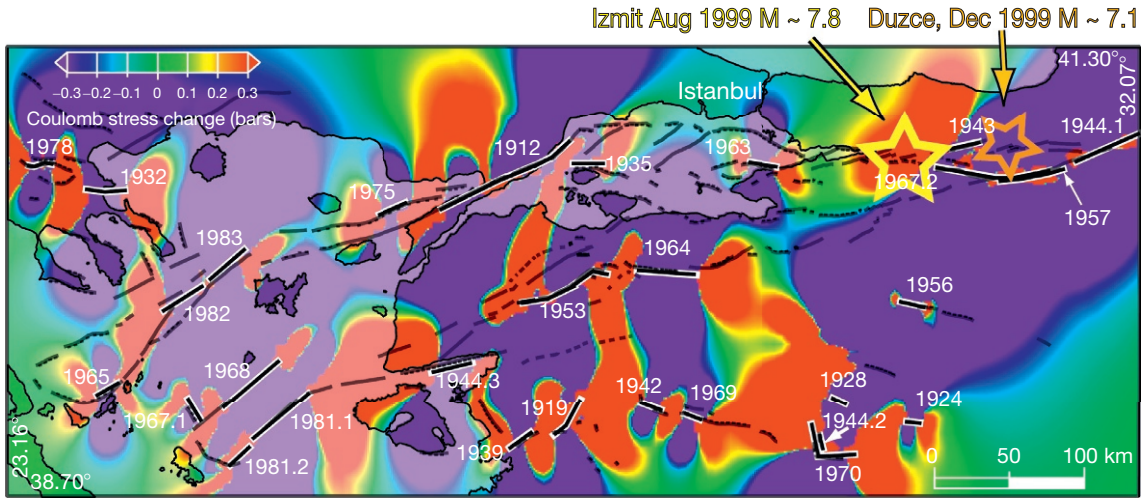
Figure 11 (a) Tectonic map of western Turkey and North Aegean Sea and distribution of seismicity between 1900 and 1996. (b) Fault plane solutions and active faults in the area (modified from Nalbant SS, Hubert A, and King GCP (1998) Stress coupling between earthquakes in northwest Turkey and the North Aegean Sea. *Journal of Geophysical Research* 103: 24469–24486). The orientation of the regional stress field used in the study is shown. Short solid lines indicate the orientation of the principal compressive stress. This is also the strike of optimally orientated normal faulting. The directions (two conjugate) for optimum strike-slip faulting are shown by lighter lines. (c) The regions and events studied by Nalbant et al. (1998), Stein et al. (1997), and Nalbant et al. (2002).



(a)



(b)



(c)

Figure 12 Coulomb stress changes on optimally oriented strike-slip and normal faults caused by earthquakes in subsequent time windows. The causative faults are shown in black lines enclosed with a white outline. Stress loaded faults are in red enclosed with a black outline. The faults in stress shadows are indicated in blue enclosed with a black outline. The black segments show those faults that are loaded by less than ±0.1 bar. Solid circles indicate aftershocks of some of the events, while open circles show background activity. (a) Stress changes caused by the events in the time window 1912–44 (1912, 1919, 1924, 1928, 1932, 1935, 1939, 1942, 1943, and 1944.1). (b) Stress changes caused by the events in the time window 1912–67 (the previous events plus those in 1944.2, 1944.3, 1953, 1956, 1957, 1963, 1964, 1965, 1967.1, and 1967.2). (c) Stress changes caused by the events in the time window 1912–96 (all the previous events plus those in 1968, 1969, 1970, 1975, 1978, 1981.1, 1981.2, 1982, and 1983). Four regions both showed increased Coulomb stress and included faults showing clear Holocene activity, part of the western Sea of Marmara, the western part of the Biga peninsula, the Saros Gulf, and the Izmit Bay region where the August 1999 (M_s 7.8) earthquake occurred.

are unlikely to experience an earthquake. Interaction studies of this sort are always limited by a lack of knowledge of events that preceded the time period studied. There are few cases in which a study can extend over a greater time period than the return periods of the earthquakes. This is the case of the study discussed in Section 4.10.5.1.

The 1999 Izmit earthquake occurred in one of four regions (see figure caption) where known active faults and increased Coulomb stress coincided. When the Izmit earthquake struck, Coulomb stress on faults in the eastern Marmara Sea increased by between 1 and 5 bar as well as by about 1 bar to the east of the Izmit rupture zone (Hubert-Ferrari et al., 2000). On 12 November 1999, a second M 7.2 earthquake (Duzce) occurred in this eastern region (see Figure 11). It is important to note that this region was in a stress shadow prior to the Izmit earthquake that changed its status. Other examples of this sort are discussed by Nalbant et al. (1998). Thus, while stress shadows are in general less likely to host future earthquakes, their status can change if an event occurs in adjacent regions. The present situation in the Sea of Marmara region is discussed in Section 4.08 where the effects of long-term loading are also included. A ~100 km stretch of fault south of Istanbul can now be identified as being likely to host a future earthquake.

The success of the Izmit and Duzce predictions together with the prediction of a M 8.7 event following the 26 December 2004 M 9.3 Great Sumatra Earthquake (McCloskey et al., 2005) provides three examples of Coulomb methods forecasting future events and forms an important part of the evidence establishing regions of enhanced Coulomb stress in places hosting known active faults as powerful approach to establishing medium-term seismic hazard for a region.

4.10.3.7 Interactions Between Very Large Earthquakes

Most of the earthquakes discussed earlier had lengths (L) that were not much greater than their down-dip width (W). Under these conditions, an earthquake on one segment of fault can significantly modify the stress on all of a nearby segment. The distance to which the stress field extends, however, depends on the smallest fault dimension; thus, when the L–W ratio becomes very large, the stress change affects only a small part of an adjacent fault segment at most (Figure 13) and thus the simple stress triggering described earlier is less likely to occur. Only if the epicenter where the next event initiates is close to the boundaries of slip of the previous earthquake is triggering unambiguous. For the North Anatolian fault sequence, six events had L–W ratios >10 and initiated close to the termination of a previous rupture. One event in 1943 initiated 230 km from the end of the rupture termination of the 1939 event and has not been regarded as being triggered by static stress changes (Stein et al., 1997). Because of the reduced role of stress interaction for large events, other effects, including loading (Section 4.10.5.1), seem to become more important for massive earthquakes.

4.10.3.8 Is It Better to Calculate Stress Transfer to Known Faults?

Most of the early studies of both aftershocks and stress interactions between large events used the formulation presented in Section 4.10.2.3 to calculate optimum fault orientations based on a regional stress field. Where the fault planes of large events were known, this was commonly backed up by calculation of Coulomb stress change on the known fault (e.g., Nalbant and McCloskey, 2011). Some recent work has abandoned looking at changes on optimum planes considering it to be better to resolve the stress onto a target fault plane (Steacy et al., 2005; Toda et al., 2008). This is correct if an averaged fault surface represents the part of the fault where triggering occurs. Faults however are inherently complex, en echelon and multistranded on many scales and rupture may initiate on local complexities. Thus, it is unknown whether the stress change resolved on the fault plane as defined by its average surface trace or focal mechanism is more appropriate than the optimum direction. It should be recalled that stress directions are

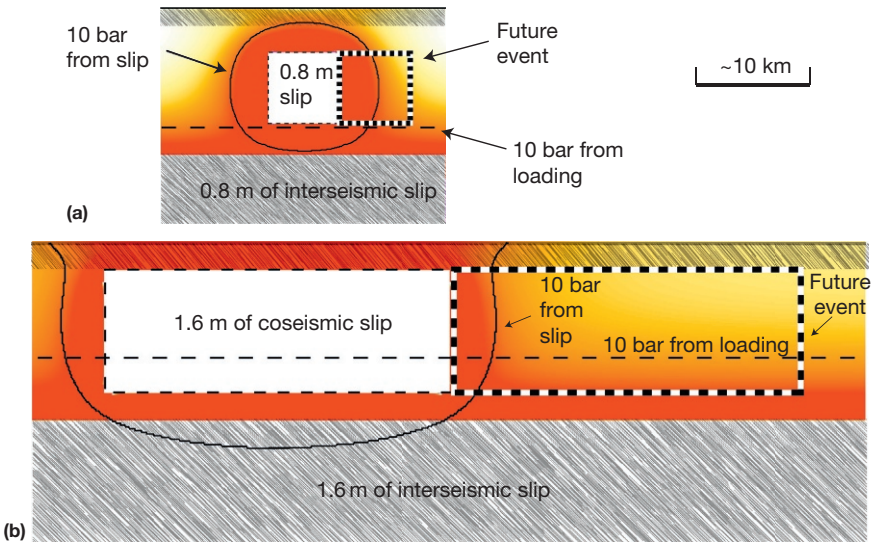


Figure 13 The relative contributions of long-term loading and seismic slip for events with length (L) to width (W) ratios of 1 and 4 for (a) and (b), respectively. The solid contour in each case indicates 10 bar of increase due to the earthquake and the dashed line 10 bar due to loading. Hatching indicates regions that slip aseismically. See also Figure 6 of Lin and Stein (2004).

derived from averages of fault directions and focal mechanisms for a region and hence may represent better those parts of a fault where slip can be triggered.

For small events and aftershocks, very large numbers of events can be examined if the optimum Coulomb condition is assumed providing a much larger data set. This could be seen as an advantage. However, since a certain knowledge of the stress system is never available, most authors have assumed that more than one mechanism is possible. An overall compressional or extensional regional stress field may allow some events to be strike-slip and not just prevailing dip-slip mechanisms. As a result, the maximum Coulomb stress for the two possible fields is chosen, for example, selecting the largest Coulomb stress increase for either strike-slip or reverse fault events. This enlarges the positive Coulomb lobes; thus, although some extra can be captured, better correlations may not be real.

Some of the errors associated with the study of aftershocks may be the failure to account for the stress changes resulting from the aftershocks themselves. Thus, the stress field from the main shock alone is inadequate to represent the stress field as an aftershock sequence evolves. The aftershock sequence can release a substantial fraction of the magnitude of the main event (Marsden and Nalbant, 2005). Clearly, the stress changes for the Big Bear aftershock could not be ignored when examining the seismicity following the Landers earthquake. However, most aftershocks are not localized in small clusters, and as the event sizes reduce, their area of influence diminishes rapidly. Consequently, except for some of the larger aftershocks, this should do no more than blur the boundaries of the main event Coulomb distribution. It certainly cannot account for events that clearly fall in regions where Coulomb stress is reduced by the main event.

As noted earlier for a number of cases, aftershock focal mechanisms are available, and a number of workers have carried out direct calculations and examined each of the two possible nodal planes and carried out various statistical tests. Perhaps surprisingly, the results have not proved to be much better than the earlier studies (see references in Steacy et al., 2005). In such tests, some aftershock sequences have worked quite well such as Landers, Superstition Hills, and Kobe, while others, notably Loma Prieta, have not. Later discussion may provide some of the reasons for the lack of improvement brought by approaches considering transfer on known faults rather than optimum directions and why studies of events such as Loma Prieta have not proved successful.

4.10.4 Introducing Time into the Failure Criterion

The results discussed so far assume that earthquakes can be represented by dislocation surfaces, that faults obey a Coulomb friction failure condition, that the lithosphere can be treated as a homogeneous, elastic half-space, and that the loading system consists of a simple stress field. The examples provided, plus many others in the literature, support the notion that stress changes due to motion on one fault can result in motion on a second. Despite examples demonstrating that other effects must also intervene, stress interaction must be an important element in any attempt to explain the most prominent feature of global earthquake statistics. Earthquake pairs are clustered in space and time with the occurrence of one event increasing the probability of a second, with the probability decaying with distance and time from the first event (Kagan and Jackson, 1991). In the development so far, time has been absent and will now be considered in subsequent sections where more complex models are adopted. Since the failure criteria on faults are time-dependent, the lithosphere is inhomogeneous, and loading conditions are not simple stress fields, attempts to account for such effects introduce a wide range of possible models.

4.10.4.1 Time-Dependent Rock Failure

Failure in a rock sample does not occur at the peak stress, but occurs in a postpeak region when the system is mechanically weakening faster than the loading system can respond. Scholz (1990) explained that the mechanical breakdown associated with this weakening requires finite displacement before failure can occur. Because this breakdown or nucleation is quasistatic, it requires a finite time to occur. It is this *kinematic* feature of failure that introduces time into failure criteria. For rock samples, this is associated with the growth and interaction of defects with the largest defects being the most important (e.g., Ashby and Sammis, 1990). The effect is to be expected both for new failure in rock and for new failure on an existing fault surface. In the lithosphere, there is no way to establish a priori scale at which the breakdown process occurs; however, for large earthquake faults, this appears to involve large volumes of rock (Manighetti et al., 2004).

Since earthquakes are commonly regarded to result from new slip on existing surfaces, the behavior of such surfaces has been carefully studied in the laboratory, giving rise to 'rate- and state-dependent friction models.' These have been widely adopted and are effective in explaining many observations. While the mathematical formulation undoubtedly captures basic features of the underlying mechanical processes, it is important to remember that failure in the lithosphere could involve different process from a saw cut in granite.

4.10.4.2 Rate- and State-Dependent Friction Equations

Many formulations of rate- and state-dependent friction law have been proposed (e.g., Scholz, 1998 and references therein). In general, these constitutive laws can be described by two coupled equations. The first one relates the sliding resistance τ to the slip velocity V and the state variables θ_i; it is usually named the governing equation,

$$\tau = F[V, \theta_i, A, B, \mu_0, \sigma_n]$$

where A and B are two positive parameters that depend on the material properties, temperature, and pressure. μ_0 is a reference friction value and σ_n is the normal stress. The second law provides the time evolution of the state variable and it is called the evolution equation. For a single-state variable, this becomes

$$\frac{\partial \theta}{\partial t} = G[V, L, \theta, B]$$

The state variable θ provides a memory for the sliding surface. A widely used formulation for these laws includes a single-state variable:

$$\tau = \mu_0 \sigma_n + A \ln \frac{V}{V^*} + B \ln \frac{\theta}{\theta^*}$$

$$\frac{\partial \theta}{\partial t} = 1 - \frac{\theta}{L} V$$

These equations have been proposed to model earthquake after effects and they have been used to predict aftershocks rates (e.g., Dieterich, 1994), to study stress shadowing (e.g., Harris and Simpson, 1996) as well as stress triggering (e.g., Gomberg, et al., 1998). Subevent triggering can occur over several hours, days, or weeks, as observed, for instance, during the 1992 Landers earthquake. During such short triggering times, the remote tectonic load is not playing any role. For longer periods, loading must be considered. The relation between short-term effects of rate and state friction and longer-term loading is illustrated in **Figure 14** modified from Stein et al. (1997). In **Figure 14(a)**, the Coulomb stress ahead of the fault is shown to increase steadily as a result of tectonic loading at a point x until it increases abruptly because of an earthquake on an adjacent fault after which it returns to the same rate. In a simple model of repeating earthquakes, a new event occurs when a threshold is crossed. The threshold is shown to be blurred to indicate that in practice, the recurrence time is irregular and has to be expressed in statistical terms – a mean and standard deviation, for example. Thus, at any given time, the probability of an earthquake in any time period depends on where the stress on the fault lies in the earthquake cycle. Shortly after an earthquake, the probability of an event is low, and as the mean recurrence time is approached, the probability of a second event being triggered increases. The simple result of the stress step is to apparently advance the proximity to the mean recurrence time and thus increase the probability that an event will occur.

The transient effects predicted by rate and state friction behave differently (**Figure 14(b)**). It is assumed that nucleation processes on and around the fault are promoted by the coseismic stress step. It thus increases the chance to trigger the fault rupture immediately after the earthquake but rapidly decays following Omori's aftershock curve. This then results in a much-greater increase of probability than the simple recurrence time model (Stein, 1999).

Numerous studies have adopted this approach. However, a study that uses the rate and state model to predict both the rates of small earthquakes and the probabilities of larger events has been published by Toda et al., 2005 accompanied by an animation that follows the evolution of seismicity in Southern California between 1986 and 2003. The study provides an articulate description of the use of rate and state friction (http://sicarius.wr.usgs.gov/landers).

4.10.5 Loading Models and Lithospheric Properties

4.10.5.1 Elastic Half-Space Loading and Stress Relaxation Models

In the introduction, two categories of lithospheric models were described: one with faults or localized shear zones extending to

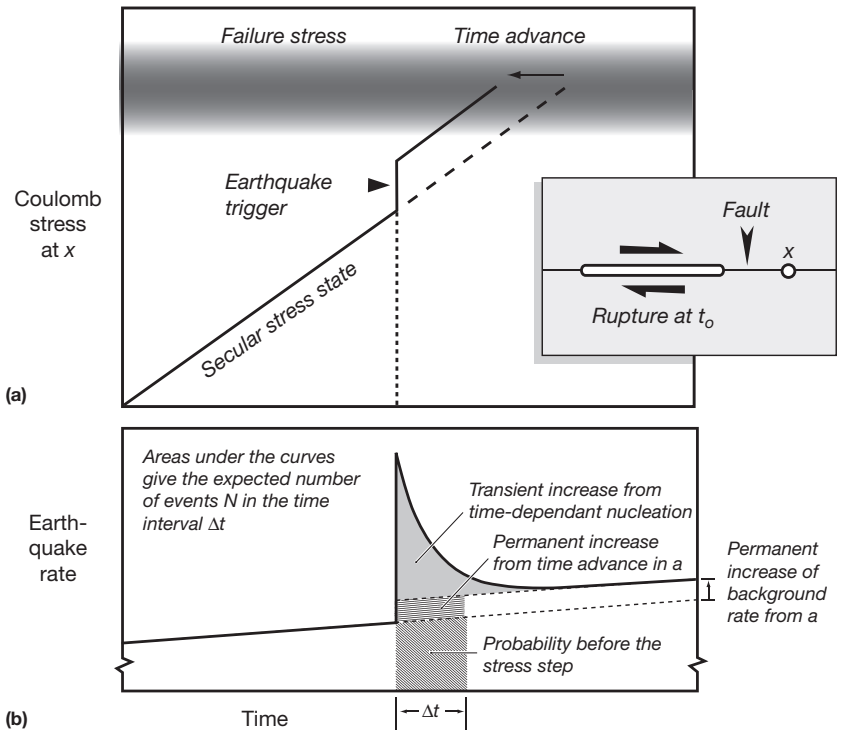


Figure 14 Schematic illustration of a sudden stress increase on a fault at point x near to an earthquake rupture at time t_0 (inset). (a) The stress increase advances the time to the next rupture. If the stress was closer to failure at the time of the stress step, then the stress step could have triggered the event immediately. (b) The associated short-term (transient) and long-term (permanent) increase in the rate of earthquake occurrence. This can also be interpreted as an increase of probability of a large event. Modified from Stein R, Barka AA, and Dieterich JH (1997) Progressive failure on the North Anatolian fault since 1939 by earthquake stress triggering. *Geophysical Journal International* 128: 594–604.

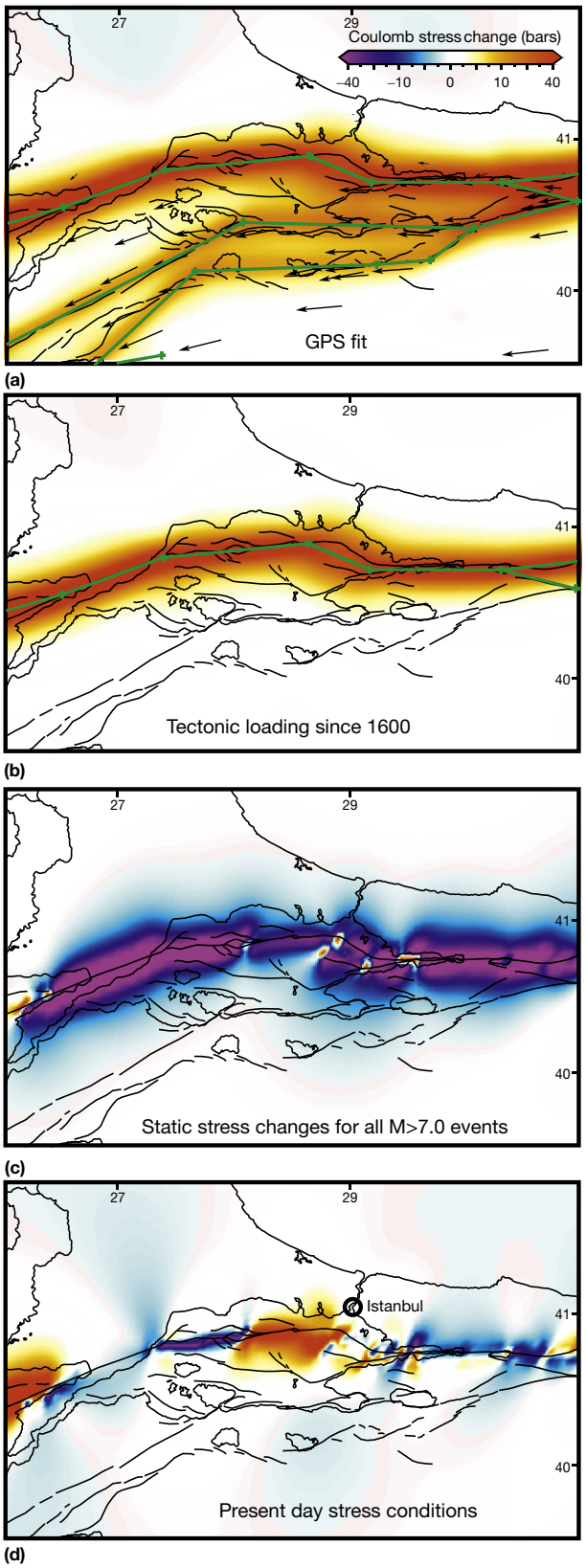


Figure 15 Accumulated loading and stress loading for the Marmara Sea region from 1600 to 2000. (a) Loading rate based by fitting GPS vectors for all branches of the Anatolian Fault around the Sea of Marmara. The panel only shows a part of region studied to determine the rates (Flerit et al., 2003). The inset scale bar corresponds to a time period of 400 years and is the same for all subsequent panels. The GPS vectors are shown as arrows and green lines identify the slipping faults beneath the locking depth. (b) The loading for the northern branch alone for 400 years. (c) Stress changes resulting from all earthquakes >M 7.0 since 1600. Smaller events are of little significance. (d) The combined effect (summed) of loading and stress changes due to events. The stressed region south and west of Istanbul is expected to host a future ~M 7 earthquake.

the asthenosphere and the second with viscous deformation occurring in the lower crust and below. Both of the approaches with various adaptations have been introduced to explain the long-term effects of loading and shorter-term time-dependent effects following big earthquakes.

From the point of view of Coulomb calculations, the widely adopted models assume that a fault continues below seismogenic depths and between earthquakes moves at a constant rate. Where a sufficient span of GPS data is available, the model used to explain geodetic observations can be used directly to generate long-term loading. An example is shown in **Figure 15(a)–15(d)** adapted from Pondard et al. (2007) that considers loading and earthquakes since 1600 for the Sea of Marmara region. **Figure 15(a)** and **15(b)** shows the stress loading based on the rate defined by GPS measurements that are also consistent with longer-term geologic rates. **Figure 15(c)** shows the stresses due to the earthquakes alone and **Figure 15(d)** combines these with the loading (**Figure 15(b)**). An ~100 km stretch of fault in the Sea of Marmara is now stressed and represents an earthquake hazard to Istanbul. Stress change and loading models of this kind have also been adopted to provide the parameters needed to estimate the probabilistic seismic hazard discussed in Section 4.08 (Parsons, 2004).

The creeping fault at depth is assumed to have a velocity hardening behavior and can consequently be expected to slip at a greater rate following an earthquake (e.g., Perfettini and Avouac, 2004). This after-slip should introduce steadily increasing stresses following an earthquake and particularly for events with large L–W ratios should cause the stressed region to expand with time (e.g., Stein et al., 1992). The effect of this evolving stress field offers a way to explain the time evolution of aftershocks that differs from that suggested by rate and state friction (**Section 4.10.4.2**). However, viscous relaxation at depth and fluid flow processes are also candidates as discussed later and in **Sections 4.10.5.2** and **4.10.5.4**.

Stress relaxation in viscous lower crust has been proposed to explain at least some aspects of the time-dependent effects associated with aftershocks and delayed triggering of more distant large events (e.g., Chery et al., 2001; Freed and Lin, 2001; Nalbant and McCloskey, 2011; Parsons et al., 2008; Pollitz et al., 2004; Toda et al., 2008). The effect, however, is difficult to distinguish in terms of the evolving stress field to the assumption that increased post-earthquake slip occurs on a fault at depth. The approach has proved successful to explain earthquake interactions in Mongolia and the 2008 Wenchuan, China, earthquake. In particular, this approach has been adopted to explain clusters of large earthquakes on the assumption that stress progressively spreads slowly from the epicentral region. The apparent success does not demonstrate that the approach is necessarily correct. Peltzer et al. (1998) pointed

out that the evolution of the vertical component of the displacement field determined from satellite-based radar is not consistent with these models. Furthermore, the behavior of the lithosphere is generally contentious (e.g., Jackson et al., 2008).

How stress is transmitted to the platelike brittle crust through the thick low viscosity lower crust proposed by authors such as Royden et al. (2008) is not easily understood nor how significant faults can develop in these models. Studies of stress interactions in loaded 2D loaded plates discussed in the next section might suggest how fault localization can occur.

4.10.5.2 2D Platelike Models

Typically, models consider the plate to be driven at a distance by displacement boundary conditions and viscous effects are ignored. Faults or deformation zones evolve in the plate as a result of purely elastic interactions between a distribution of initial defects that grow and interact by obeying a simple failure law. Studies of this sort adopt automata modeling (e.g., Ferguson et al., 1999; Olami et al., 1992; Rundle et al., 1995). These approaches assume that the seismogenic layer is 100% seismogenic and all earthquakes are a consequence of earlier events. This assumption also underlies models that call on 'self-organized criticality' and is an underlying justification for concepts such 'epidemic-type aftershock sequence' models (e.g., Helmstetter and Sornette, 2002; Ogata, 1988, 1989; Ogata et al., 2003). These models are proposed to have a universal application (e.g., Helmstetter and Sornette, 2002; Helmstetter, et al., 2003; Ogata, 2004; Sornette and Helmstetter, 2002).

While these techniques may produce a useful statistical approach to characterizing earthquake catalogs and some insight into interaction processes, a problem lies in differences in interaction between faults with small and large L–W ratios and their coupling to loading processes resulting from stable sliding both within and beneath the seismogenic zone. This is summarized in **Figure 16**. **Figure 16(a)** shows that the stick-slip region does not come fully to the surface and even within the seismogenic region patches of stable sliding can occur. This effect can be seen in the depth distribution of slip contributed by events at different depths (**Figure 16(b)**). Near to the surface and near the base of the seismogenic layer stick-slip events contribute only a modest amount of slip (King et al., 1994a). In creeping regions such as central California, only a few percent of the slip is seismic and occurs on a few patches (**Figure 16(c)**). Events in such regions are too widely separated to interact (**Figure 16(d)**). Unlike the situation shown in **Figure 13** in the center of a noncreeping seismogenic zone where interactions are strong, it seems that complete stress coupling and stick-slip behavior can only occur in a narrow region in the center of a seismogenic zone. Since events of $M \sim 6$ completely cut the seismogenic zone, only much smaller events in a narrow depth range can be considered to be dominated by elastic interactions alone. The related problem of large earthquakes not belonging to the same fractal set as small ones is considered by Scholz (1990) and Pacheco et al. (1992).

4.10.5.3 Inhomogeneous Lithospheric Models: The Effects of Voids or Fissures

The most dramatic inhomogeneity that can be introduced in a solid is a void, which can contain volcanic magma, water, or gas. A familiar feature of volcanic or thermal regions is the earthquakes that occur very close to a volcanic vent or hot springs. The b-value of the Gutenburg–Richter relation associated with such earthquakes is commonly high – a large number of small events occur for each larger event. Expressed in terms of moment release, small events account for as much deformation as large events unlike the seismicity of nonvolcanic regions (see also **Section 4.10.5.5**). The same character is shared by the seismicity of creeping faults, and in both cases, the deformation associated with the seismicity is tiny compared with that due to nonseismic processes (see also **Figure 16**). The most straightforward explanation is that earthquakes in volcanic regions like those on creeping faults are not interacting, being separated by voids that deform aseismically. Sibson (2002) offered a hypothetical geometry for such processes based on the work of Hill (1977).

4.10.5.4 Interactions Between Volcanic Eruptions and Earthquakes

Since volcanic fissures or magma chambers can be treated using dislocation theory, the techniques described for the interaction between large earthquakes can be modified to examine interactions between volcanoes and earthquakes and vice versa. Unlike earthquake faults, however, the plumbing of volcanoes is less well understood. It is not clear whether an eruption should be regarded as resulting from an increase of magma pressure forcing magma to the surface or a reduction of normal stress allowing fissures to open and the magma to reach the surface. It is similarly not clear whether the magma is contained in quasi-spherical chamber or in fissures. The various possibilities result in different stress fields whether examining the triggering of eruptions by earthquakes or the reverse. Notwithstanding these possible ambiguities, plausible relations between static stress changes associated with earthquakes and eruptions have been proposed for Djibouti and Italy (e.g., Noir et al., 1997; Nostro, et al., 1998; Walter and Amelung, 2007). Relations between volcanoes and earthquakes including dynamic and viscoelastic effects are reviewed by Hill et al. (2002).

4.10.5.5 Fluid Movement in Nonvolcanic Regions

Substantial evidence is accumulating for the mobilization of water in the seismogenic crust following major earthquakes both close to the active fault and in a region extending for several fault lengths from the epicenter. Direct evidence is derived from water expelled at the surface by the compressional elastic rebound of normal faulting earthquakes (Muir-Wood and King, 1993). Monitored changes of river flows finds that the increased flow progressively decays over 6–10 months (**Figure 17**). Comparing the volume of water expelled to the volumetric strain calculated from a source known from seismic and geodetic data indicates that a substantial proportion of the coseismic stress is relaxed by fluid flow. The volume expelled and the time constants involved suggest that the water is held in a connected system of high-aspect-ratio cracks. The importance of fluid flow is also suggested from geodetic studies of post-earthquake relaxation (e.g., Jonsson et al., 2003; Peltzer et al., 1998). As noted earlier in these cases, the relative effects of relaxation in a viscous lower crust, post-earthquake creep on

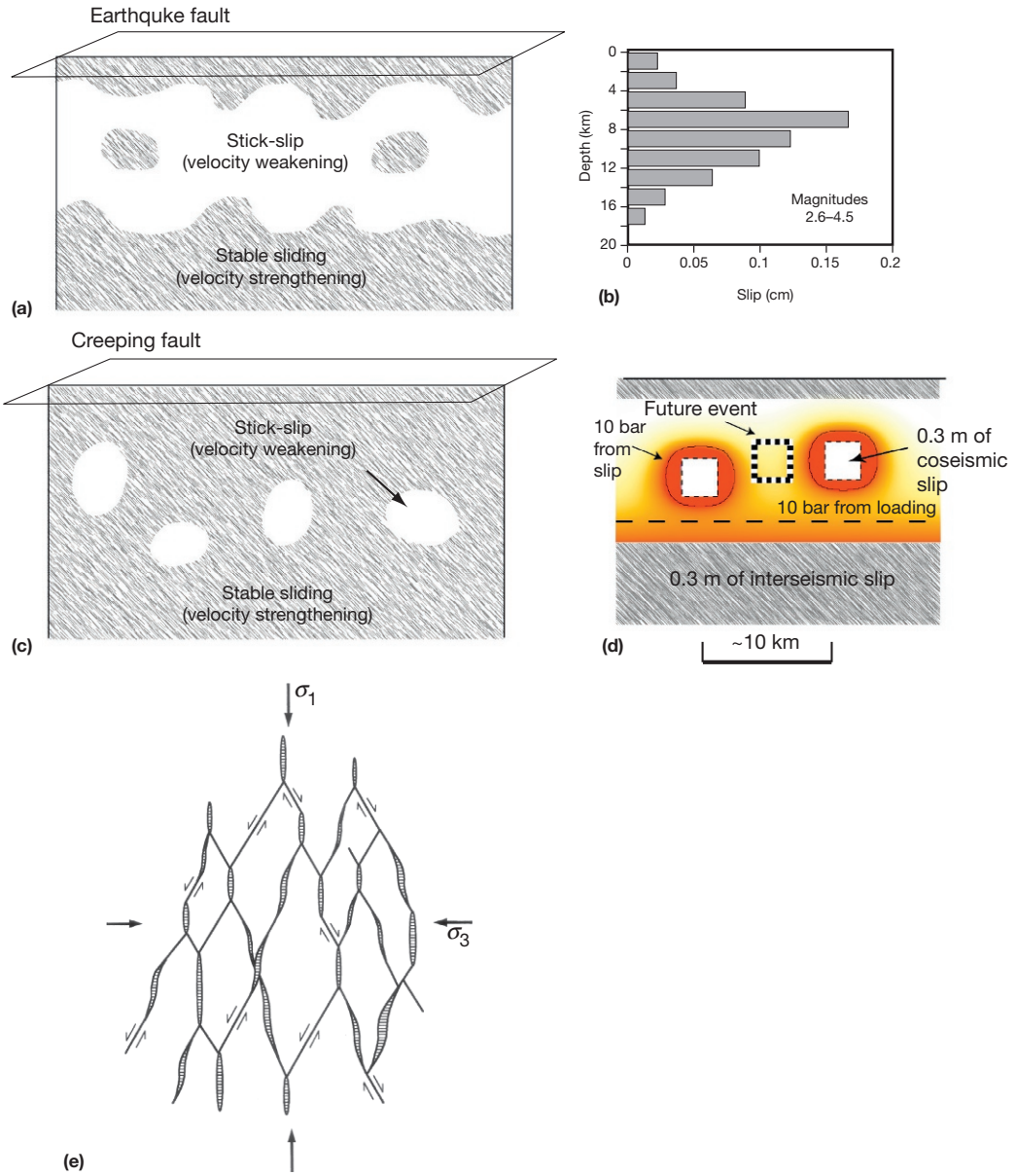


Figure 16 (a) Strike-parallel longitudinal section showing schematically regions of stick-slip and stable sliding. (b) The depth distribution of slip for events with magnitudes between 2.6 and 4.5 for the San Francisco Bay Area. The seismic slip is concentrated around 8 km. Similar results are found for other magnitude ranges (King et al., 1994a), indicating that at shallow and greater depths in the seismogenic zone, much of the deformation is aseismic as indicated in (a). A discussion of the depth of seismicity within the seismogenic zone can be found in Pacheco et al. (1992). (c) Strike-parallel longitudinal section showing schematically regions of stick-slip and stable sliding for a fault where aseismic creep predominates. (d) For a creeping fault, small stick-slip regions are too far apart to interact by stress coupling. (e) A mesh model for swarm seismicity involving interlinked minor faults and extension fractures infilled with dikes, hydrothermal veins, etc. (crosshatched), developed in relation to the principal compressive stresses, $\sigma_1 > \sigma_2 > \sigma_3$. Reproduced from Hill DP (1977) A model for earthquake swarms. *Journal of Geophysical Research* 82: 1347–1352.

a lower crustal fault and the effects of water flow are not easily separated.

The importance of the fluid flow suggested by these studies indicates that it must play some controlling role in aftershock sequences, and the time period of about 8 months is similar to the length of aftershock sequences. Two effects may control seismicity. The first is a change of the effective friction due to changes of fluid pressure (eqn [6]; e.g., Nur and Booker, 1972), thus altering the rate and distribution of seismicity predicted from Coulomb stress changes alone. While the Coulomb stress distribution may not be greatly changed, fluid pressure changes may be an important alternative to rate and state friction to explain temporal changes. Furthermore, increases of pressure could trigger events in regions where reduced Coulomb stress predicts an absence of activity.

The suggestion that water is contained in high-aspect-ratio cracks and changes in pressure over a period of several months indicate that the effective rock modulus will change over a similar period. Since water is held in cracks whose orientation reflects the prevailing stress field, the modulus changes can be

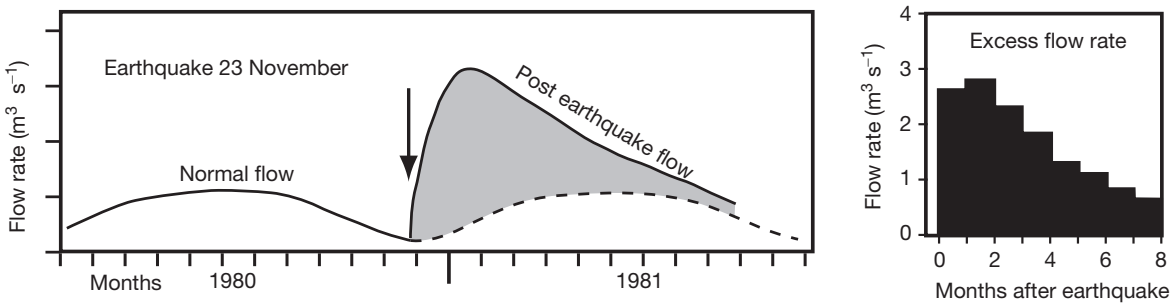


Figure 17 Excess flow for the Caposele spring resulting from compressional elastic rebound following M_s 6.9 Irpinia earthquake 1980. The form of the excess flow is typical for normal faulting events. Reproduced from Muir-Wood R and King GCP (1993) Hydrological signatures of earthquake strain. *Journal of Geophysical Research* 98: 22035–22068.

expected to be anisotropic. Such readjustments may be expected to generate stress fields and hence seismicity that are not simply related to a regional stress field or the Coulomb Stress change field following an earthquake. Several publications on the topic can be found in the review papers in the introduction. Of particular interest are recent observations of a correlation between the monsoon and midcrustal seismicity in the Himalayas (Bollinger et al., 2007) and an aftershock sequence driven by high-pressure carbon dioxide rising from depth (Miller et al., 2004).

4.10.5.6 The Distribution of Small Events Before Large Earthquakes

The aftershocks that result from a large earthquake are unambiguously due to stress increases that result from the event (**Section 4.10.3**). Over longer periods of time, there is also clear evidence for earlier large earthquakes setting up the stress conditions that control future large events. In the case of the Landers earthquake, smaller events also appear to have loaded the site of the future event (**Figure 8** and associated text). The effect of long-term loading is also important and becomes progressively more significant for events with a large L–W ratio where interactions become less important (**Section 4.10.3.7**). Since the 1980s, considerable interest has been directed to the study of accelerated moment release (AMR) (e.g., Jaumé and Sykes, 1996), where activity is found to increase over a region with dimensions several times greater than the fault length (L). A range of explanations have been offered to explain the observations. Two relate directly to the concept of stress interactions. In the first, stress triggering between small events is considered to progressively create a stress field sufficiently homogeneous over a large enough region to host a big earthquake. The process is considered to be a manifestation of self-organized critical behavior. This attractive model is unfortunately difficult to relate to the observations of the geologists, geodesists, or seismologists who observe large earthquakes repeating on readily identified faults. Some of these problems have been addressed by the conditional criticality model (Bowman and Sammis, 2004).

A second explanation attributes the increase of activity as the consequence of the stress in a wide region increasing toward failure to be the consequence of the final stages of reloading a fault. This can be thought of as the final refilling of an earthquake shadow resulting from an earlier event. The final process of loading is shown schematically in **Figure 18** (King and Bowman, 2002; Scholz, 1990), with activity progressively increasing as the stress shadow fills.

Following an earthquake, the region that must have been loaded prior to the event can be calculated in a similar way to that used to identify post-earthquake Coulomb stress increases but by slipping the fault in the opposite direction to the slip that occurred in the event. It should be appreciated that this back-slip approach (Savage, 1983) models only the large-scale stress field that drives the earthquake. More local stress fields associated with fault complexity that progressively produce topography are not modeled in this way (e.g., Perfettini and Avouac, 2004). For all the events with magnitude >6.5 in California, an acceleration of activity has been found. When compared to random catalogs, the individual probability that the AMR for each event is a physical effect is modest. However, the cumulative probability that all events show AMR as a result of random processes is <1%. Hardebeck et al. (2008) argued that it has no significance compared to their synthetic catalogs. The AMR in the far-field region of the 26 December 2004 Sumatra earthquake prior to the event (**Figure 19**) had <2% probability of being random (Mignan et al., 2006). While for Californian events with L–W ratios of ~3.5, effects other than loading can possibly be invoked to explain AMR, for the Sumatra event with an L–W ratio of ~35, loading seems the only plausible explanation for the observed pre-event increase of seismic activity.

4.10.5.7 Stress Inhomogeneity Resulting from Displacement Boundary Conditions

The foregoing sections have assumed that it is possible to define a regional stress field and steadily increasing stress due to tectonic loading. This is only perturbed by motion on known faults and it is these stress changes that trigger earthquakes. Calculations performed using the mathematics of dislocations in a homogeneous elastic half-space are assumed to provide an adequate approximation.

Prior to the success and popularity of Coulomb stress triggering models, other kinematic explanations for off-fault aftershocks had been suggested. These were based on a realization that distributed seismicity can be due to the fragmentation that results from a homogeneous material being subject to certain displacement boundary conditions (see King, 1986; King and Nabelek, 1985). While this approach has been eclipsed by the

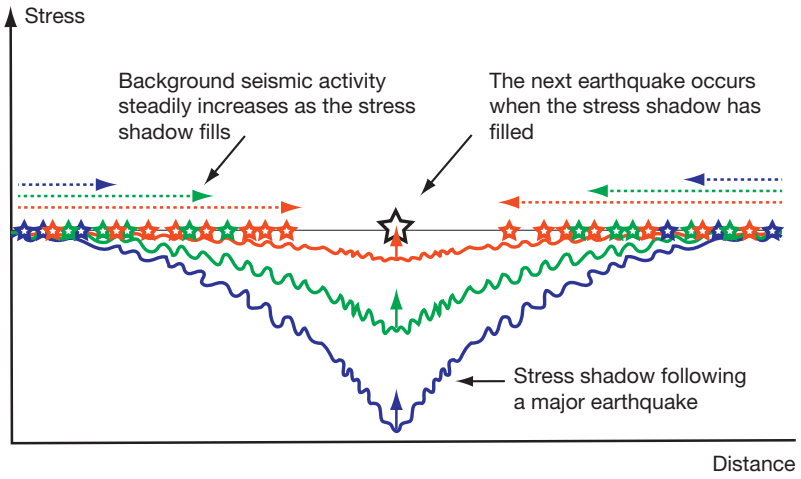

Stress
Background seismic activity steadily increases as the stress shadow fills
The next earthquake occurs when the stress shadow has filled
Stress shadow following a major earthquake
Distance

Figure 18 Following a large earthquake, a stress hole or shadow is created around the associated fault (blue line). This progressively fills (green then red lines) until the next major earthquake occurs. For the filling of the stress shadow for events with large *L–W* ratios, the effect of aseismic creep at depth predominates. As the stress shadow fills, activity increases over a wide region. Modified from Scholz C (1990) *The Mechanics of Earthquakes and Faulting*. New York: Cambridge University Press, 439 pp; King GCP and Bowman DD (2002) The evolution of regional seismicity between large earthquakes. *Journal of Geophysical Research* 107: 2354.

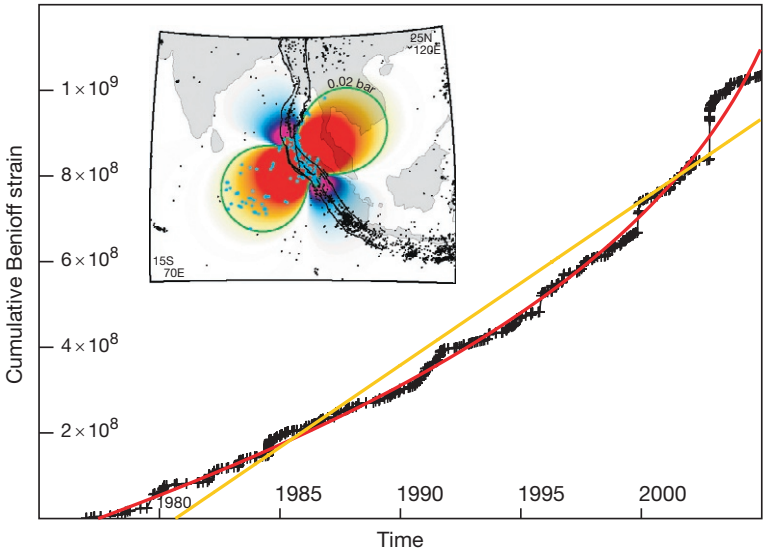


Figure 19 Accelerating moment release (AMR) for the great Sumatra earthquake of 26 December 2004. The cumulative Benioff strain and the time (in years) are, respectively, the *y*-axis and *x*-axis. The power law fit (red curve) has less than a 2% probability of occurring randomly. The prestress field calculated by the back-slip earthquake method for the high-slip part of the 26 December 2004 event is shown in the inset. The red contour (0.02 bar) outlines the area of events that provides the AMR. The fault used has four pure dip-slip segments each with a dip of 11°, a down-dip width of 200 km. Reproduced from Mignan A, King GCP, Bowman D, Lacassin R, and Dmowska R (2006) Seismic activity in the Sumatra–Java region prior to the December 26, 2004 ($M_w = 9.0–9.3$) and March 28, 2005 ($M_w = 8.7$) earthquakes. *Earth and Planetary Science Letters* 217: 339–408.

stress transfer models described in earlier chapters, it remains important for developing a fuller appreciation of how faults interact. The underlying concepts are summarized in Figure 20. Figure 20(a) shows that for simple shear displacement boundary conditions, motion can be accommodated by strike-slip faults with a single strike direction. For simple contraction or extension (Figure 20(b)), antithetic dip-slip faults can also, in principle, accommodate long-term motion provided the faults completely cut from the surface to a depth at which stress is relaxed by flow. In these two cases, the faults are kinematically stable because they release the boundary conditions by accumulating finite displacement without changing geometry. But no other boundary conditions can be accommodated in such a simple way (Figure 20(c)). A system of faults with various orientations must come into existence even though the material was previously homogeneous and unfractured. Where these faults meet kinematic incompatibility requires further multiple faulting.

The behavior of this faulting can be examined using a generic model of three strike-slip boundaries meeting at a triple junction, a system identified to be always unstable (McKenzie and Morgan, 1969). In a plate tectonic context, one or more of

the boundaries must allow the creation or destruction of plate area (i.e., must be ridges or trenches) for a triple junction to be stable and evolve in a straightforward way. In general, deforming continental lithosphere lacks such sources and sinks so the evolution of a three-fault triple junction evolution is complex (illustrated in Figure 21). In Figure 21(a)), finite fault slip requires a volume change at the junction. In the lithosphere, however, high confining pressure prevents large volume changes and deformation must accumulate off the main faults (Figure 21(b)). Although small deformation can be accommodated elastically, substantial deformation must be accommodated by fracture (Figure 21(c) and 21(d)). The predicted

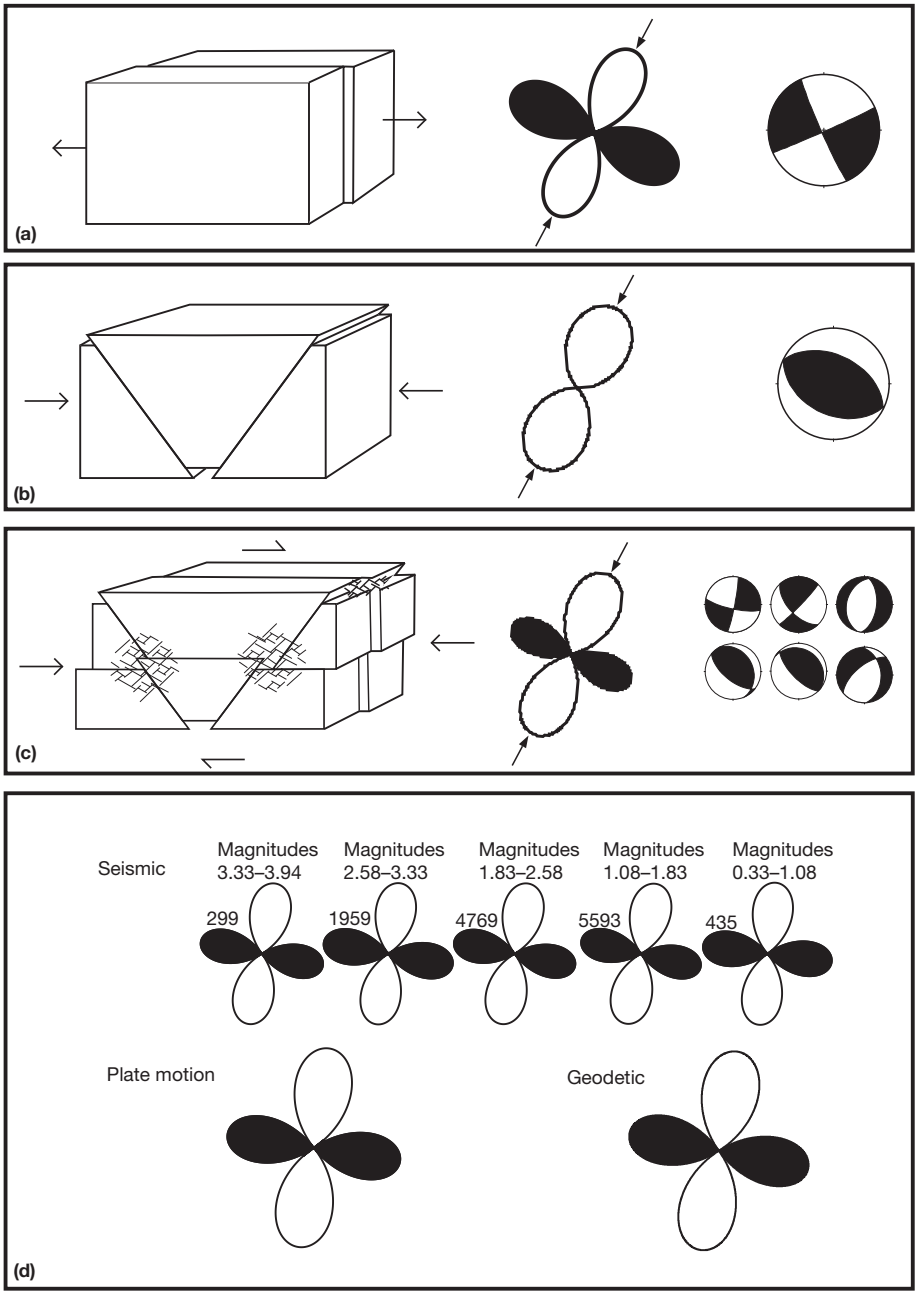


Figure 20 The kinematics of finite strain. (a) Simple shear displacement boundary can be released on a single fault. The associated strain rosette and focal mechanism are shown. (b) Simple contraction can be relieved (under certain circumstances) by simple dip-slip faulting. The associated strain rosette and focal mechanism are shown. (c) More complex displacement boundary conditions require multidirection faulting indicated schematically. The strain rosette can only result from numerous earthquakes with different mechanisms operating together. The mechanisms shown are typical of the heterogeneous aftershocks for the Loma Prieta earthquake. (d) The seismic strain can be represented by examining earthquakes in magnitude windows each of which includes earthquakes that change average fault length by a factor of 2. The distribution of finite strain released by small events has the same form as that predicted from plate motion and determined geodetically. Parts of the figure are adapted from Amelung F and King GCP (1997) Large scale tectonic deformation inferred from small earthquakes. *Nature* 386: 702–705.

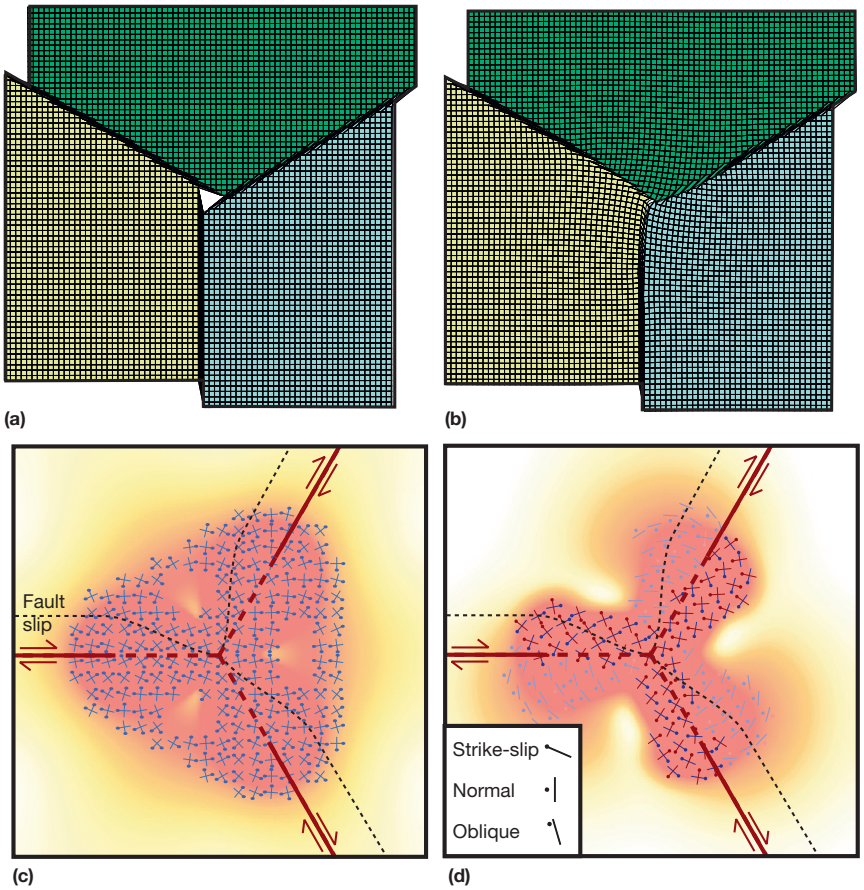


Figure 21 Junctions between strike-slip faults are kinematically unstable (McKenzie and Morgan, 1969). They are however observed at the Earth's surface (e.g., junctions between the North and East Anatolian Faults in Turkey). The deformation associated with such junctions can be accommodated in two ways. (a) A void is created. The faults separate three undeforming blocks. (b) In the Earth's lithosphere, however, confining pressure prevents large volume changes. At the junction, slip must come to zero forcing deformation to be accommodated off the main faults. (c) For a plane strain model (2D), the shear strain (in red) induced around the junction by motion on the main faults is shown. Dashed curves illustrate the slip distribution on the main faults. Slip comes to zero at the junction between the three faults. Symbols represent the predicted slip directions (see **Figure 7**; Devès et al., 2011 and inset below in **Figure 24**). Whereas mechanisms are all strike-slip, they present different directions preventing localization to occur on long distances. Deformation can only be accommodated in a distributed manner, which can occur, for instance, by motion on many small faults (or shear zones) with different directions. (d) For a plane-stress model with a compressional out of plane principle strain, the strain distribution is changed and some mechanisms are normal faulting, but again deformation has to remain distributed; localization cannot develop on long distances. Figure reproduced from Devès M, King GCP, Klinger Y, and Amotz A (2011) Localised and distributed deformation in the lithosphere: Modelling the Dead Sea region in 3 dimensions. *Earth and Planetary Science Letters* 308: 172.

directions of fracture are shown in these figures and can be seen to change from place to place around the junction. No fault can grow to a substantial length; therefore, deformation is taken up by small-scale faulting occurring in multiple directions. In two dimensions, three directions must exist to accommodate the strain, and in three dimensions, five directions must exist (King, 1983). Small-scale faults also interact to create areas of kinematic instability resulting in multiple faulting at ever-smaller scale. At the scale of the main strike-slip faults, however, off-fault deformation appears to be uniformly distributed around the junction. The term 'process zone' has been used to refer to these areas (e.g., Devès et al., 2011; Scholz, 2002).

While the triple junction of **Figure 21** provides a simple example, other geometric complexities such as fault bends, offsets, or intersections are all places of kinematic instability requiring some deformation to occur in a distributed manner (King and Nabelek, 1985). **Figure 20(a)–20(c)** shows that three deformation systems can have the same principle axis of shortening, but in the case of **Figure 21(c)**, several different focal mechanisms must operate to generate the observed azimuthal strain distribution (strain rosette). The earlier assumption that the principle stress axes uniquely define the orientation of fault planes is not correct and the multifault deformation described requires that the concept of a temporally constant regional stress field should be abandoned. Stresses must be both inhomogeneous in space and evolve with time. The observation of complex fault mechanisms is a consequence of the kinematics of finite deformation and not as

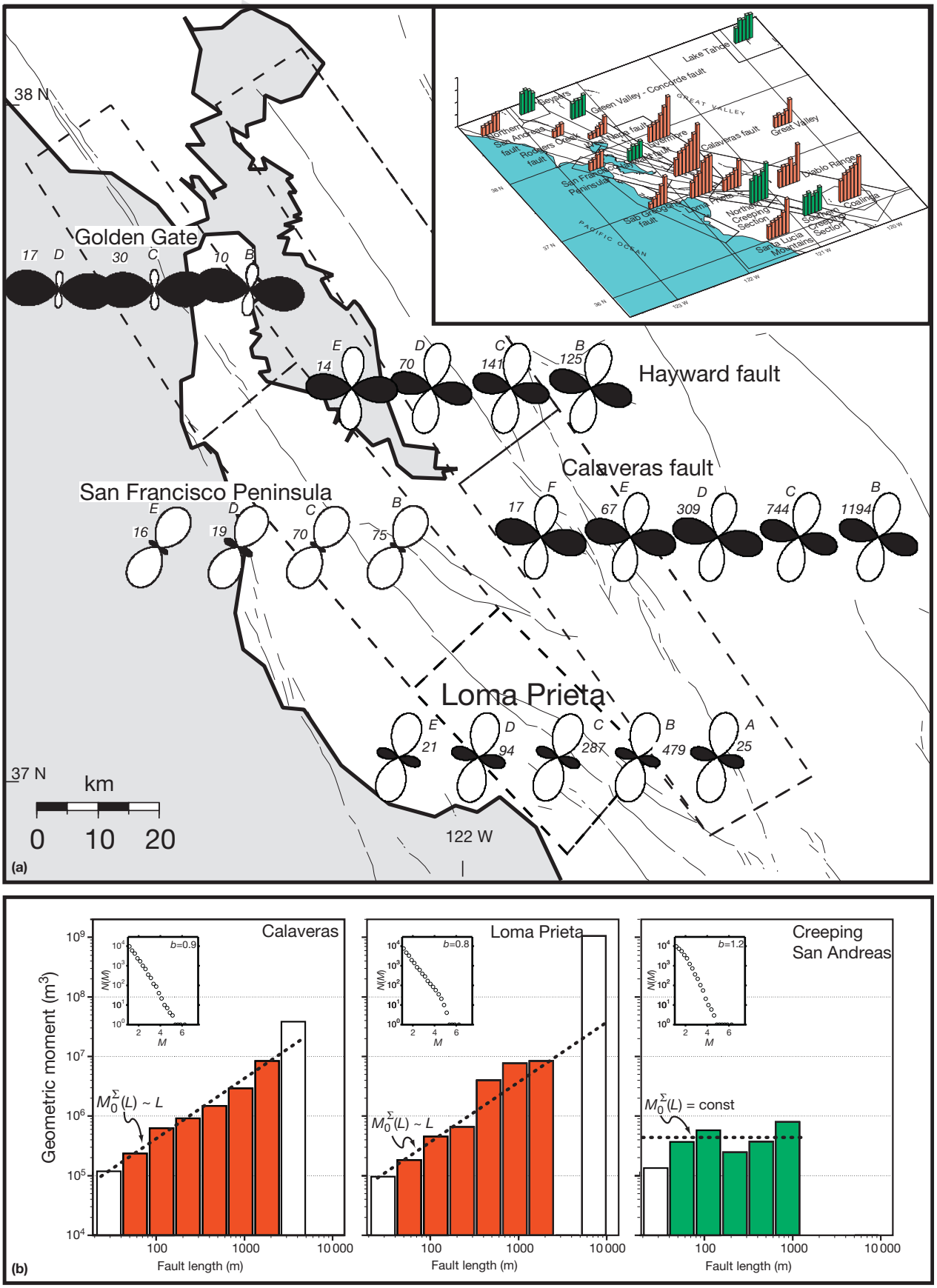


Figure 22 Earthquake kinematics for the San Francisco Bay Area. The strain rosettes are created by separating events into bins with average fault dimensions differing by factors of 2 (as described in the caption to **Figure 17**). Although the overall finite strain summed for the whole region is the same as the applied boundary conditions (**Figure 17(d)**), there are substantial local variations. (b) The scaling relations for the summed geometric moment (potency) in the bins described earlier, for example, regions. For noncreeping regions, the moment release doubles for each factor of

(continued)

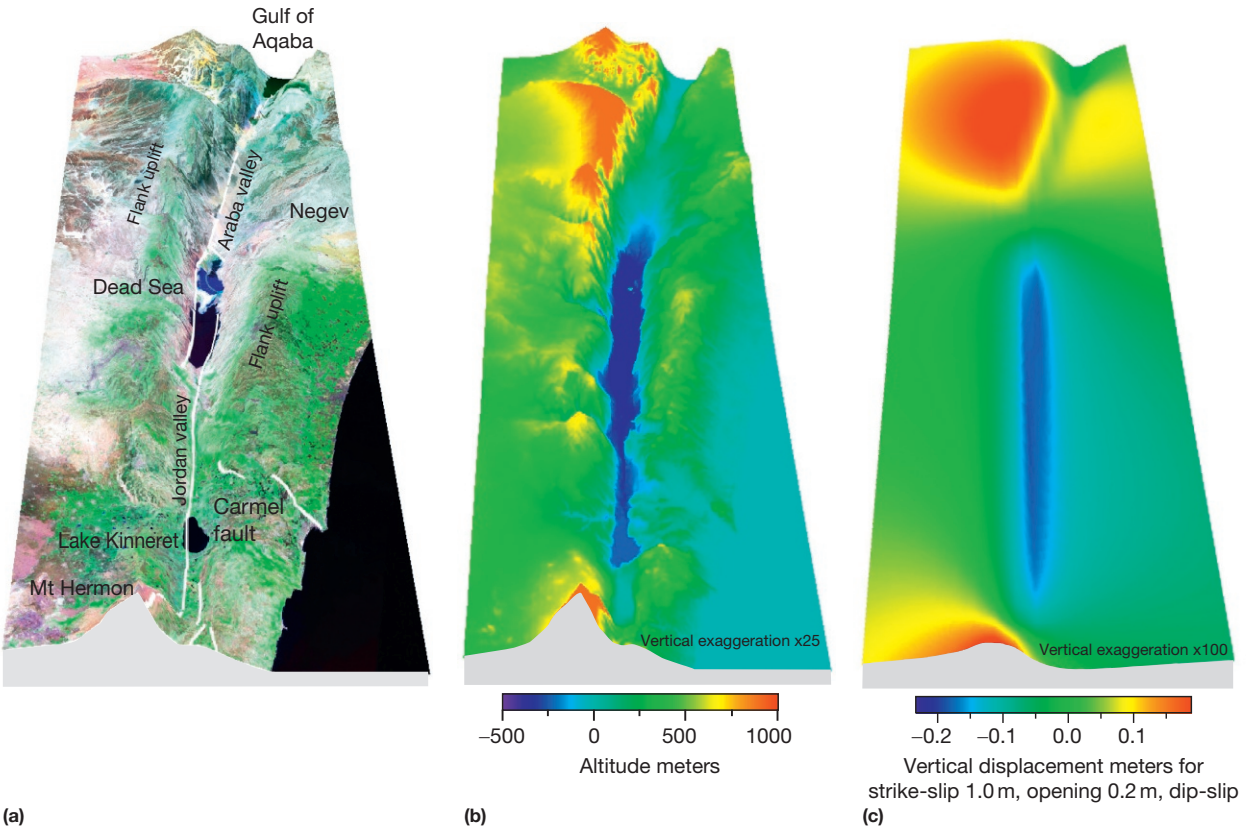


Figure 23 Topographic representations of the Dead Sea valleys (a) shows a Landsat Thematic Mapper image (band 2, red; 4, green; and 7, blue) draped overexaggerated digital elevation data (SRTM 3) to give a 3D effect. The vertical exaggeration is 25. The most prominent feature is the 11.5° change of valley strike occurring at the Dead Sea. While the valley is a clear feature, there are no clear normal faults marking the valley flanks (unlike Carmel Fault) such as in true rifts like Corinth or East Africa. (b) Shows a 3D image of the topography (SRTM3 data) colored by altitude. (c) Shows uplift and subsidence rates predicted by the model. Figure reproduced from Devès M, King GCP, Klinger Y, and Amotz A (2011) Localised and distributed deformation in the lithosphere: Modelling the Dead Sea region in 3 dimensions. *Earth and Planetary Science Letters* 308: 172.

commonly assumed the result of low friction on preexisting faults of varying orientation. Whatever the value of fault friction, multiple mechanisms are the only way to accommodate complex boundary conditions.

Although stresses associated with multimechanism faulting are inhomogeneous, the strain that is relieved by a collection of mechanisms can be simple (Amelung and King, 1997a,b). **Figure 20(d)** shows strain rosettes for the whole San Francisco Bay Area. Except for the largest magnitude ranges where there are insufficient events, the deformation due to all the small events summed in magnitude ranges is the same as that predicted by the overall plate motion and geodesy. The forms of the rosettes for subregions within the San Francisco Bay region are shown in **Figure 22**. These are consistent with geologic observations of the ongoing tectonics. For the Calaveras and Hayward faults, the strain rosettes can be explained by simple strike-slip faulting. Along the San Francisco peninsular and the Loma Prieta region, this is not the case; the strain conditions also require contractional faulting. The subsiding Golden Gate region where the great 1906 earthquake is thought to have initiated is associated with substantial extension. This event could have started on a fault with a substantial normal component of slip and triggered by stresses very different from those that would be appropriate if the fault is represented by the average strike and mechanism of the San Andreas. This expands further on the remarks in **Section 4.10.3.8** that suggest that resolving Coulomb stresses onto a plane determined from field measurements may not be appropriate.

Devès et al. (2011) have expanded on the kinematic approach modeling the Dead Sea valley system as an example. They use a distribution of dislocations to model the displacement field by matching the topography (**Figure 23**). This allows the strain field to be established that in turn allows the fault mechanisms to be predicted. **Figure 24** shows that in the middle of the valley, colinear strike-slip mechanisms are predicted consistent with the kinematically stable direction followed by the main strike-slip Dead Sea fault. For the flanks of the valley both

Figure 22—cont'd 2 increase of fault dimensions. For creeping faults and volcanic or thermal spring regions, the moment release is the same for all bins. Insets show the equivalent *b*-values. The inset to (a) shows the observed scaling for regions in North California. The figures are simplified from Amelung and King (1997a,b). The rosettes are labeled with letters between A and F that correspond to magnitude bins A, 1.08–2.58; B, 2.58–3.33; C, 3.33–3.94; D, 3.94–4.54; E, 4.54–5.14; F, 5.14–5.75.

at the surface and at depth, stable faults cannot form requiring this deformation to be accommodated by multifault fragmentation. The modeling allows previously unexplained features of the valley to be accounted for. Although widely described as the Dead Sea Rift, the valley is not bounded by classic normal faults but by down warps composed of multiple and irregular faulting. The focal mechanisms also show a range of mechanisms. Both of these features can be explained by the distributed (process zone) deformation (Devès et al., 2011). There are other areas such as the western Basin and Range in the United States where multiple faulting occurs, but which have not been studied in the way described.

The purpose of this chapter is to emphasize that while stress and stress transfer models can be powerful, they can capture only a part of how faults interact and suggest why (as noted in Section 4.10.3.8) in complex regions such as Loma Prieta Coulomb modeling of the aftershock sequence has not been successful.

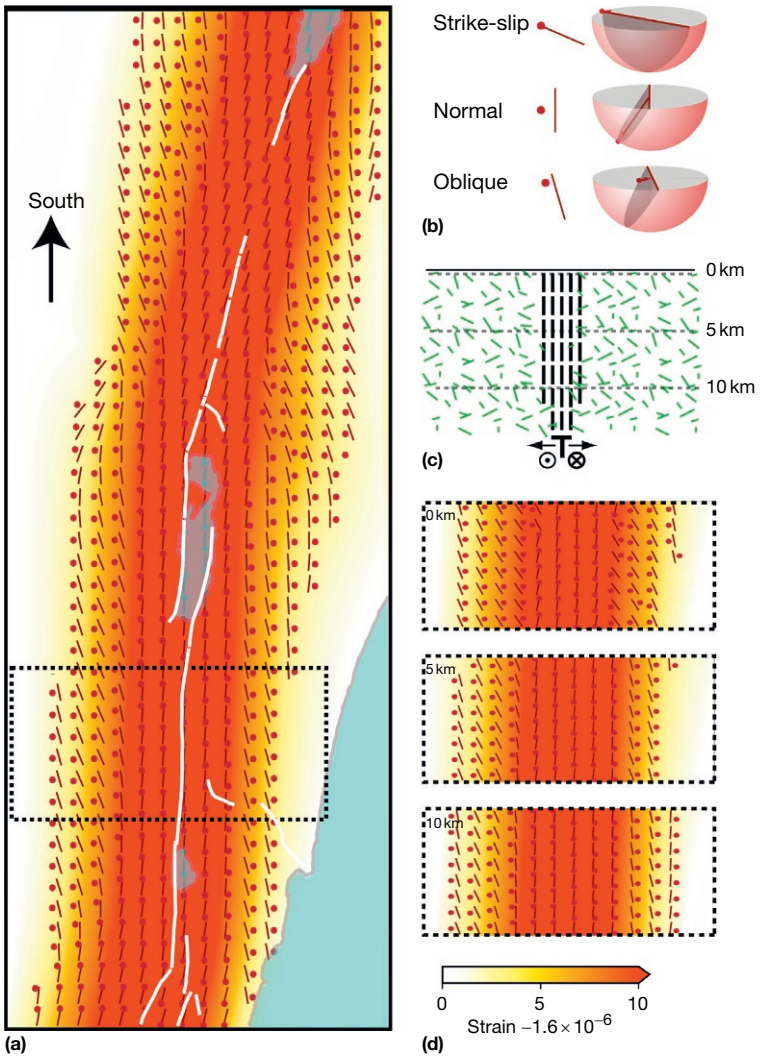


Figure 24 Predicted slip directions for the Jordan Rift. A regional vertical strain of -1.6×10^{-6} and a horizontal strain of 3.2×10^{-6} at $-40°$ E and of 0.0 at 50° W are applied. (a) Shows predicted directions for strike-slip, normal and oblique faults with left-lateral components at a depth of 5 km. The shading indicates strain (=geometric moment per unit volume). Note that only one of two possible dip directions is shown for normal faults. An inset, bottom right, shows the scale. Note that across the width of the rift the color scale is saturated. Only the strike-slip mechanisms near the center of the rift allow the creation of long kinematically stable faults parallel to the valley axis. These are consistent with the main active strike-slip faults. Many mechanisms on the flanks are normal with an extension direction ~50° E. Others have oblique mechanisms with varying ratios of normal to strike-slip components. Normal faulting parallel to the valley is not predicted. A change in the distribution of dislocation sources used to model deformation can slightly change the predicted slip directions but will not affect our main conclusions – flank deformation cannot localize on single throughgoing faults. (b) Shows the interpretation of the mechanisms shown in (a), (c) and (d). (b) Shows a schematic cross section indicating where strike-slip faulting can form (vertical lines) and where faulting cannot localize (green hatching) and indicates the depths at which (a) and (d) are calculated. (d) Shows predicted directions for different depths. Strike-slip mechanisms are predicted at all depth along the valley axis. Away from the axis, many mechanisms become oblique and are different at different depths. Normal faults parallel to the valley cannot develop. Since the predicted mechanisms are not kinematically stable, other faulting mechanisms must develop resulting in process zones of distributed deformation that creates the observed form of the valleys. Figure reproduced from Devès M, King GCP, Klinger Y, and Amotz A (2011) Localised and distributed deformation in the lithosphere: Modelling the Dead Sea region in 3 dimensions. *Earth and Planetary Science Letters* 308: 172.

4.10.6 Dynamic Triggering

This section concerns static stress triggering – its successes and failures. When discussing the latter, the possibility of events being triggered by large stresses associated with the passage of seismic waves must be considered. There is no doubt that dynamic triggering occurs; indeed, the propagation of rupture on a fault plane is a process of continuous triggering. At distances too great to result from static stress changes, events occur that are clearly related to the distant earthquake. The events in Alaska associated with 26 December 2004 Sumatra earthquake are a stunning example (West et al., 2005).

The dynamic triggering of events, however, cannot be explained by a failure occurring at a simple stress threshold as discussed in **Section 4.10.2**. Except near to the epicenter, the peak dynamic stresses associated with propagating seismic waves are much greater than those associated with the static stress change. Threshold triggering would result in every very big earthquake, triggering events all over the world. Because failure takes time (**Section 4.10.4**), this does not occur. Even large transient stresses are less important than sustained stresses.

If delayed failure suppresses dynamic triggering, why does it occur? At great distances, triggering is commonly associated with fluid processes, either water or magma, suggesting that special conditions are needed for it to occur (Hill et al., 2002). These are discussed in **Chapter 4.11**. Until the processes involved are better understood, events that appear in the negative lobes of static Coulomb stress change must be considered to be a possible consequence of dynamic stresses.

4.10.7 Conclusions

Relating the occurrence of one earthquake to another using changes of Coulomb stress is clearly successful. The distribution of aftershocks can be effectively explained in many cases. Interactions between large events have also been demonstrated and used to successfully predict destructive earthquakes (in Turkey and Sumatra). A problem with this approach is the modest length of most historical catalogs. This information can be recovered by geologic methods (fault geomorphology and trenching combined with steadily improved methods of dating), which will in due course greatly improve our understanding of earthquake hazard. Of equal importance, the concept of stress shadows can be used to suggest where destructive events are less likely to occur (**Section 4.10.3.6**). However, where these stress shadows fall along major faults, they do not last indefinitely. Adjacent loading over time restores the region to failure prior to the next substantial event. This is associated with increasing background activity in the region. Where it can be identified (**Section 4.10.5.7**), this provides a further approach to identifying the location of future damaging events. These results are of major social importance.

Stress interactions do not occur instantaneously – time is involved. The most successful approach to quantifying this invokes the concept of rate and state friction (**Section 4.10.4.2**). Following an abrupt stress increase, the rate at which small events occur increases. A combination of Coulomb stress change and rate and state friction can explain both the distribution of aftershocks while the Omori law describes their rate of decay. The same approach can be adapted to predict an enhanced earthquake probability of a second event following a first. Such earthquake probability forecasts form the most effective way of communicating risk to government authorities, emergency services, and the general public, but it will be many years before the approach can be scientifically validated. Models that consider steadily increased loading as a result of viscous relaxation of the lower crust also aim to predict possibly times to future events (Nalbant and McCloskey, 2011; Parsons et al., 2008).

Despite the successes outlined, in many cases, earthquakes occur where they are not predicted by Coulomb stress changes alone. Many possible explanations have been discussed. Many draw on the fact that lithosphere is not a homogeneous elastic half-space, faults are not fully described by dislocation surfaces, and stress boundary conditions do not capture important features of finite deformation. Dynamic triggering can also explain features that cannot be explained by static stress models.

The problem with advancing our understanding of how earthquakes interact lies in finding data that can be interpreted unambiguously. At present, there are more models possible than data to distinguish between them. However, improvements in the data needed to understand lithospheric mechanics should come to our aid. For example, recent advances in our ability to map faults with the aid of satellite images and determine their rates with techniques such as cosmogenic dating. Space geodesy, both GPS and InSAR, is also placing constraints on the behavior of the lithosphere. In due course, these will allow a more of how and with what earthquakes interact. While we wait for the future, we are lucky that simple ideas work quite well and can certainly save lives.

Acknowledgment

The authors thank Renata Dmowska and Shinji Toda for helpful remarks. This is IPGP contribution number 2214 and INSU contribution number 396.

References

Aki K and Richards P (1980) *Quantitative Seismology*. San Francisco, CA: W. H. Freeman and Co.

Amelung F and King GCP (1997a) Large scale tectonic deformation inferred from small earthquakes. *Nature* 386: 702–705.

Amelung F and King GCP (1997b) Earthquake scaling laws for creeping and non-creeping faults. *Earth and Planetary Science Letters* 24: 507–510.

Ashby MF and Sammis CG (1990) The damage mechanics of brittle solids in compression. *Pure and Applied Geophysics* 133: 489–521.

Bettinelli P, Avouac JP, Flouzat M, et al. (2008) Seasonal variations of seismicity and geodetic strain in the Himalaya induced by surface hydrology. *Earth and Planetary Science Letters* 266(3): 332–344.

Bowman DD and Sammis C (2004) Intermittent criticality and the Gutenberg–Richter distribution. *Pure and Applied Geophysics* 161: 1945–1956.

Chery J, Carretier S, and Ritz JF (2001) Postseismic stress transfer explains time clustering of large earthquakes in Mongolia. *Earth and Planetary Science Letters* 194(1): 277–286.

Cocco M and Rice JR (2002) Pore pressure and poroelasticity effects in Coulomb stress analysis of earthquake interactions. *Journal of Geophysical Research* 107: 2030.

Devès M, King GCP, Klinger Y, and Amotz A (2011) Localised and distributed deformation in the lithosphere: Modelling the Dead Sea region in 3 dimensions. *Earth and Planetary Science Letters* 308: 172.

Devès MH, Tait S, King GCP, and Grandin R (in press) Strain heating in process zones; implications for metamorphism and partial melting of the litohsphere. *Earth and Planetary Science Letters*.

Dieterich J (1994) A constitutive law for the rate of earthquake production and its application to earthquake clustering. *Journal of Geophysical Research* 99: 2601–2618.

Ferguson CD, Klein W, and Rundle JB (1999) Spinodals, scaling, and ergodicity in a threshold model with long-range stress transfer. *Physical Review E* 60(2): 1359.

Flerit F, Armijo R, King GCP, Meyer B, and Barka A (2003) Slip partitioning in the Sea of Marmara pull-apart determined from GPS velocity vectors. *Geophysical Journal International* 154(1): 1–7.

Freed AM (2005) Earthquake triggering by static, dynamic, and postseismic stress transfer. *Annual Review of Earth and Planetary Sciences* 33: 335–367.

Freed AM and Lin J (2001) Delayed triggering of the 1999 Hector Mine earthquake by viscoelastic stress transfer. *Nature* 411: 180–183.

Gomberg J, Beeler NM, Blanpied ML, and Bodin P (1998) Earthquake triggering by transient and static deformations. *Journal of Geophysical Research: Solid Earth (1978–2012)* 103(B10): 24411–24426.

Hardebeck JL, Feltzer KR, and Michael AJ (2008) Improved tests reveal that the accelerating moment release hypothesis is statistically insignificant. *Journal of Geophysical Research* 113: B08310. http://dx.doi.org/10.1029/2007JB005410.

Harris RA (1998) Introduction to special section: Stress triggers, stress shadows, and implications for seismic hazard. *Journal of Geophysical Research* 103: 24347–24358.

Harris RA and Simpson RW (1996) In the shadow of 1857 – The effect of the great Ft. Tejon earthquake on subsequent earthquakes in southern California. *Geophysical Research Letters* 23: 229–232.

Helmstetter A and Sornette D (2002) Sub-critical and super-critical regimes in epidemic models of earthquake aftershocks. *Journal of Geophysical Research* 107: 2237.

Helmstetter A, Sornette D, and Grasso JR (2003) Mainshocks are aftershocks of conditional foreshocks: How do foreshock statistical properties emerge from aftershock laws. *Journal of Geophysical Research: Solid Earth (1978–2012)* 108(B1).

Hill DP (1977) A model for earthquake swarms. *Journal of Geophysical Research* 82: 1347–1352.

Hill DP, Pollitz F, and Newhall C (2002) Earthquake–volcano interactions. *Physics Today* 55: 41–47.

Hubert-Ferrari A, et al. (2000) Seismic hazard in the Marmara Sea following the 17 August 1999 Izmit earthquake. *Nature* 404: 269–272.

Jackson JA, McKenzie DP, Priestley KF, and Emmerson B (2008) New views on the structure and rheology of the lithosphere. *Journal of the Geological Society* 165(2): 453–465.

Jaumé SC and Sykes LR (1996) Evolution of moderate seismicity in the San Francisco Bay region, 1850 to 1993: Seismicity changes related to the occurrence of large and great earthquakes. *Journal of Geophysical Research* 101: 765–790.

Jonsson S, Segall P, Pedersen R, and Björnsson G (2003) Post-earthquake ground movements correlated to pore-pressure transients. *Nature* 424(6945): 179–183.

Kagan YY and Jackson DD (1991) Long-term earthquake clustering. *Geophysical Journal International* 104: 117–133.

King GCP (1983) The accommodation of large strains in the upper lithosphere of the earth and other solids by self-similar fault system: The geometrical origin of b-value. *Pure and Applied Geophysics* 121: 761–815.

King GCP (1986) Speculations on the geometry of the initiation and termination of earthquake rupture and the evolution of morphology and geological structures. *Pure and Applied Geophysics* 124: 567–585.

King GCP and Bowman DD (2002) The evolution of regional seismicity between large earthquakes. *Journal of Geophysical Research* 107: 2354.

King GCP and Cocco M (2000) Fault interaction by elastic stress changes: New clues from earthquake sequences. *Advances in Geophysics* 44: 1–38, Academic Press.

King GC and Nabelek J (1985) The role of fault bends in faults in the initiation and termination of earthquake rupture. *Science* 283: 984–987.

King GCP, Oppenheimer D, and Amelung F (1994a) Block versus continuum deformation in the western United States. *Earth and Planetary Science Letters* 128(3): 55–64.

King GCP, Stein RS, and Lin J (1994b) Static stress changes and the triggering of earthquakes. *Bulletin of the Seismological Society of America* 84(3): 935–953.

Lin J and Stein RS (2004) Stress triggering in thrust and subduction earthquakes and stress interaction between the southern San Andreas and nearby thrust and strike-slip faults. *Journal of Geophysical Research: Solid Earth (1978–2012)* 109(B2).

Manighetti I, King GCP, and Sammis CG (2004) The role of off-fault damage in the evolution of normal faults. *Earth and Planetary Science Letters* 217(3): 399–408.

Marsden D and Nalbant S (2005) Methods of measuring seismicity rate changes: A review of how the M 7.3 Landers earthquake affected the aftershock sequence of the M 6.1 Joshua Tree earthquake. *Pure and Applied Geophysics* 162: 1151–1185.

McCloskey J, Nalbant SS, and Steacy S (2005) Indonesian earthquake: Earthquake risk from co-seismic stress. *Nature* 434(7031): 291.

McKenzie DP and Morgan WJ (1969) Evolution of triple junctions. *Nature* 224: 125.

Mignan A, King GCP, Bowman D, Lacassin R, and Dmowska R (2006) Seismic activity in the Sumatra–Java region prior to the December 26, 2004 (Mw = 9.0–9.3) and March 28, 2005 (Mw = 8.7) earthquakes. *Earth and Planetary Science Letters* 217: 339–408.

Miller SA, Collettini C, Chiaraluce L, Cocco M, Barchi M, and Kaus BJ (2004) Aftershocks driven by a high-pressure CO_2 source at depth. *Nature* 427(6976): 724–727.

Mindlin R (1936) Force at a point in the interior of a semi-infinite solid. *Physics* 195–202.

Muir-Wood R and King GCP (1993) Hydrological signatures of earthquake strain. *Journal of Geophysical Research* 98: 22035–22068.

Nalbant SS, Hubert A, and King GCP (1998) Stress coupling between earthquakes in northwest Turkey and the North Aegean Sea. *Journal of Geophysical Research* 103: 24469–24486.

Nalbant SS and McCloskey J (2011) Stress evolution before and after the 2008 Wenchuan, China earthquake. *Earth and Planetary Science Letters* 307(2011): 222–232.

Nalbant S, McCloskey J, Steacy S, and Barka AA (2002) Stress accumulation and increased seismic risk in eastern Turkey. *Earth and Planetary Science Letters* 195: 291–298.

Noir J, Jacques E, Bekri S, Adler PM, Tapponnier P, and King GCP (1997) Fluid flow triggered migration of events in the 1989 Dobi earthquake sequence of Central Afar. *Geophysical Research Letters* 24: 2335–2338.

Nostro C, Stein RS, Cocco M, Belardinelli ME, and Marzocchi W (1998) Two-way coupling between Vesuvius eruptions and southern Apennine earthquakes, Italy, by elastic stress transfer. *Journal of Geophysical Research* 103: 24487–24504.

Nur A and Booker JR (1972) Aftershocks caused by pore fluid flow. *Science* 175: 885.

Ogata Y (1988) Statistical models for earthquake occurrence and residual analysis for point processes. *Journal of the American Statistical Association* 83: 9–27.

Ogata Y (1989) Statistical model for standard seismicity and detection of anomalies by residual analysis. *Tectonophysics* 169: 159–174.

Ogata Y (2004) Space–time model for regional seismicity and detection of crustal stress changes. *Journal of Geophysical Research* 109: B06308.

Ogata Y, Jones LM, and Toda S (2003) When and where the aftershock activity was depressed: Contrasting decay patterns of the proximate large earthquakes in southern California. *Journal of Geophysical Research* 108(B6): 2318.

Okada Y (1985) Surface deformation to shear and tensile faults in a half space. *Bulletin of the Seismological Society of America* 75: 1135–1154.

Olami Z, Feder HJS, and Christensen K (1992) Self-organized criticality in a continuous, nonconservative cellular automaton modeling earthquakes. *Physical Review Letters* 68(8): 1244.

Pacheco JF, Scholz CH, and Sykes LR (1992) Changes in frequency–size relationship from small to large earthquakes. *Nature* 355(6355): 71–73.

Parsons T (2004) Recalculated probability of $M > 7$ earthquakes beneath the Sea of Marmara, Turkey. *Journal of Geophysical Research* 109: B05304. http://dx.doi.org/10.1029/2003JB002667.

Parsons P, Ji C, and Kirby E (2008) Stress changes from the 2008 Wenchuan earthquake and increased hazard in the Sichuan basin. *Nature* 454: 509–510. http://dx.doi.org/10.1038/nature07177.

Peltzer G, Rosen P, Rogez F, and Hudnut K (1998) Poro-elastic rebound along the Landers 1992 earthquake surface rupture. *Journal of Geophysical Research* 103: 30131–30145.

Perfettini H and Avouac JP (2004) Stress transfer and strain rate variations during the seismic cycle. *Journal of Geophysical Research* 109: B06402.

Pollitz F, Bakun WH, and Nyst M (2004) A physical model for strain accumulation in the San Francisco Bay region: Stress evolution since 1838. *Journal of Geophysical Research* 109: B11408. http://dx.doi.org/10.1029/2004JB003003.

Pondard N, Armijo R, King GCP, Meyer B, and Flerit F (2007) Fault interaction in the Sea of Marmara pull-apart (North Anatolian Fault): Earthquake clustering and propagating sequences. *Geophysical Journal International* 171: 1185–1197. http://dx.doi.org/10.1111/j.1365-246X.2007.03580.

Royden LH, Burchfiel BC, and van der Hilst RD (2008) The geological evolution of the Tibetan Plateau. *Science* 321: 1054. http://dx.doi.org/10.1126/science.1155371.

Rundle JB, Klein W, Gross S, and Turcotte DL (1995) Boltzmann fluctuations in numerical simulations of nonequilibrium lattice threshold solids. *Physical Review Letters* 75: 1658–1661.

Savage JC (1983) A dislocation model of strain accumulation and release at a subduction zone. *Journal of Geophysical Research* 105: 3095–3102.

Scholz C (1990) *The Mechanics of Earthquakes and Faulting*. New York: Cambridge University Press, 439 pp.

Scholz CH (1998) Earthquakes and friction laws. *Nature* 391: 37–42.

Scholz CH (2002) *The Mechanics of Earthquakes and Faulting*. New York: Cambridge university press.

Sibson RH (2002) Geology of the crustal earthquake source, edited, pp. 455–470. International Association for Seismology and Physics of the Earth's Interior, Committee on Education.

Sornette D and Helmstetter A (2002) On the occurrence of finite-time-singularities in epidemic models of rupture, earthquakes and starquakes. *Physical Review Letters* 89: 158501.

Steacy S, Gomberg J, and Cocco M (2005) Introduction to special section: Stress transfer, earthquake triggering, and time-dependent seismic hazard. *Journal of Geophysical Research: Solid Earth (1978–2012)* 110(B5).

Stein RS (1999) The role of stress transfer in earthquake occurrence. *Nature* 402: 605–609.

Stein R, Barka AA, and Dieterich JH (1997) Progressive failure on the North Anatolian fault since 1939 by earthquake stress triggering. *Geophysical Journal International* 128: 594–604.

Stein RS, King GCP, and Jian L (1992) Change in failure stress on the southern San Andreas fault system caused by the 1992 Magnitude.7.4 Landers earthquake. *Science* 258: 1328–1332.

Suppe J (1985) *Principals of Structural Geology*. Englewood Cliffs, NJ: Prentice-Hall, 537 pp.

Toda S, Lin J, Meghraoui M, and Stein RS (2008) 12 May 2008 M = 7.9 Wenchuan, China, earthquake calculated to increase failure stress and seismicity rate on three major fault systems. *Geophysical Research Letters* 35: L17305. http://dx.doi.org/10.1029/2008GL034903.

Toda S, Ross S, Stein RS, Gregory C, Beroza G, and David Marsan D (2012) Aftershocks halted by static stress shadows. *Nature Geoscience* 5: 410–413. http://dx.doi.org/10.1038/ngeo1465, Published online 6 May.

Toda S, Stein RS, Richards-Dinger K, and Serkan B (2005) Forecasting the evolution of seismicity in southern California: Animations built on earthquake stress transfer. *Journal of Geophysical Research* 110(B5): B05S16.

Volterra V (1907) Sur l'equilibre des corps elastiques multiplement connexes. *Annales scientifiques de l'École Normale Supérieure. Paris* 401–517.

Walter TR and Amelung F (2007) Volcanic eruptions following M ≥ 9 megathrust earthquakes: Implications for the Sumatra–Andaman volcanoes. *Geology* 35: 539–542.

West M, Sánchez JJ, and McNutt SR (2005) Periodically triggered seismicity at Mount Wrangell, Alaska, after the Sumatra earthquake. *Science* 308(5725): 1144–1146.

This page intentionally left blank

4.11 Dynamic Triggering

DP Hill, US Geological Survey, Volcano Science Center, Menlo Park, CA, USA
SG Prejean, US Geological Survey, Alaska Volcano Observatory, Anchorage, AK, USA

© 2015 Elsevier B.V. All rights reserved.

4.11.1	**Introduction**	274
4.11.1.1	Stress Transfer Modes	274
4.11.1.2	Static and Quasistatic Stress Triggering	274
4.11.1.3	Dynamic Stress Triggering	275
4.11.1.4	Triggered or Coincidental?	276
4.11.2	**Early Inferences on Dynamic Triggering and Preinstrumental Examples**	276
4.11.3	**Instrumental Evidence for Dynamic Triggering**	277
4.11.3.1	Widespread Dynamic Triggering in the Seismogenic Crust	278
4.11.3.2	Dynamic Triggering of Volcanic/Geothermal and Hydrologic Systems	282
4.11.3.3	Dynamic Triggering of Tectonic (Nonvolcanic) Tremor in the Transitional Domain	284
4.11.4	**Triggered Response Characteristics**	285
4.11.4.1	The Global Search for a Triggering Threshold	285
4.11.4.2	Aftershocks and Stress Triggering at Near-to-Intermediate Distance	287
4.11.4.3	Tectonic Setting and Source Radiation Patterns	288
4.11.4.4	Onsets, Delay Times, Durations, and Repeat Intervals	290
4.11.4.5	Triggering and Modulation by Solid Earth Tides	291
4.11.4.6	Earthquake Hazard Implications of Dynamic Triggering	291
4.11.5	**Proposed Models**	292
4.11.5.1	Triggering by Frictional Failure	292
4.11.5.1.1	Coulomb failure under dynamic stresses	292
4.11.5.1.2	Nonlinear friction	296
4.11.5.2	Triggering Through Excitation of Crustal Fluids	296
4.11.5.2.1	Hydrous fluid transport: changes in permeability and pore pressure	297
4.11.5.2.2	Magmatic fluids	297
4.11.6	**Challenges for the Future**	298
4.11.7	**Conclusions**	299
Acknowledgments		299
References		299

Abbreviations

CFF	Coulomb failure function
ETAS	Statistical model for epidemic-type aftershock sequences
ETS	Episodic tremor and slip

Symbols

σ_n^*	Rock matrix stress component
c	Constant in Omori's law (eqn [1])
C	Cohesive strength
C_T	Tensile strength
g	Gravitational acceleration
G_i, R_i	($i = 1, 2, \ldots$) Love and Rayleigh waves circling the globe i times
k	Constant in Omori's law (eqn [1])
K_i	Constant values for CFF; $i = 1, 2, 3, \ldots$ (Figure 8(a))
L	Semimajor axis of a dynamic stress orbit
M	Local magnitude
M_0	Seismic moment
Ms	Surface-wave magnitude
M_w	Moment magnitude
$n(t)$	Number of earthquakes per unit time
p	Constant in Omori's law (eqn [1])
P	Pore pressure
$P(\gamma)$	Triggering potential as a function of incidence angle
R, $R(t)$	Radius in Mohr's circle ½ $(\sigma_1 - \sigma_3)$: see Figure 8
t	Time
t_0	Arrival time of dynamic stresses (seismic waves)
t_1	Onset of triggered response
t_2	Termination of dynamic stresses
Vp, Vs	P- and S-wave phase velocities

 http://dx.doi.org/10.1016/B978-0-444-53802-4.00078-6

ΔCFF	Static change in CFF
Δt	Delay time ($t_1 - t_0$)
Δ	Epicentral distance (km); Rayleigh wave dilatational stress
T_p	Peak dynamic stress
T	Traction acting on a fault plane at angle in Mohr's circle diagram (Figure 9)
β	Beta statistic
δCFF(t)	Dynamic change in CFF
ε_p	Peak dynamic strain
γ	Seismic-wave incidence angle with respect to fault strike
$\mu_d(t)$	Dynamic coefficient of friction
μ_s	Static coefficient of friction
θ_c	Critical angle for failure on an optimally oriented fault
$\theta(t)$	Angle between the least principal stress, σ_3, and a fault plane
ρ	Density
σ_1, σ_3	Greatest and least principal stress components
σ_m	Mean stress, ½ ($\sigma_1 + \sigma_3$)
σ_n	Effective normal stress component ($\sigma_n^* - P$)
σ_z	Vertical principal stress
τ, τ_s, τ_d	Shear stress components: τ_s strike parallel; τ_d dip parallel
$\lvert\delta\text{CFF}\rvert_m$	Absolute value of maximum δCFF(t) (see Figure 8(b))

4.11.1 Introduction

One of the outstanding questions in earthquake seismology involves the nature of short-term processes that ultimately trigger slip on a given fault. The Earth's crust is pervasively laced with faults, and multiple lines of evidence indicate that the brittle crust is in some sense critically stressed and in a state of incipient failure nearly everywhere (Zoback and Zoback, 2002). Because long-wavelength stresses associated with the relative motion of major tectonic plates accumulate relatively slowly (rates of $\sim 10^{-5}$ MPa year^{-1} or less), regional stresses may remain incrementally below the frictional strength of faults (typically on the order of ~10–100 MPa) for periods of decades to centuries. Insight into the triggering question thus depends on understanding the influence of short-term, short-wavelength fluctuations in both the stress field and the fault strength, such that the local stress state temporarily exceeds the local failure threshold leading to slip nucleation and an earthquake.

Sources of short- to mid-term stress fluctuations in the crust are many. Those most likely to be significant at seismogenic depths include (1) other earthquakes in the crust (Das and Scholz, 1981; Freed, 2005; Harris, 1998; Steacy et al., 2005; **Chapter 4.10**), (2) magmatic intrusions (Hill et al., 2002; Manga and Brodsky, 2006; Savage and Clark, 1982), (3) anthropogenic activities such as reservoir filling, mining, and fluid injection or withdrawal (McGarr et al., 2002), (4) solid Earth tides (Cochran et al., 2004; Thomas et al., 2009, 2012), (5) seasonal meteorologic factors such as snow loading and groundwater recharge (Christiansen et al., 2005; Miller, 2008), and (6) long-term static and quasistatic stressing by ocean loading (Luttrell and Sandwell, 2010) and glacial rebound (Wu and Johnston, 2000). Local fluctuations in fault strength may result from changes in fluid pore pressure within a fault zone (Lockner and Beeler, 2002), the nonlinear response of fault zone friction to small perturbations in local stresses (Dieterich, 1979, **Chapter 4.04**, this volume; Johnson and Jia, 2005), and subcritical crack growth (Atkinson, 1984).

In this chapter, we are concerned with slip nucleation triggered either directly or indirectly by dynamic stress. Our principal focus will be on the evidence for dynamic stresses in the form of seismic waves triggering earthquakes in the seismogenic crust (the brittle domain) and tectonic tremor (nonvolcanic tremor) in the brittle–plastic transition at depths of 20–40 km along major plate-boundary faults. We include evidence for tidal triggering as a low-frequency, low-amplitude reference point for the spectral range of oscillatory stressing with the potential for dynamic triggering. Because many sites displaying remotely triggered seismicity are volcanic and geothermal areas, our scope includes the response of magmatic, hydrothermal, and hydrologic systems to dynamic stresses as potential sources for locally triggered seismicity. We conclude with a perspective on key challenges to be met in advancing understanding of dynamic triggering and its implications for active crustal processes. Also see Prejean and Hill (2009).

4.11.1.1 Stress Transfer Modes

Earthquake–earthquake interactions in the form of mainshock–aftershock sequences have been recognized since the late 1800s and the emergence of seismology as a quantitative science (Omori, 1894). With the global expansion of high-quality seismic networks coupled with the capabilities for resolving detailed spatial–temporal variations in deformation patterns afforded by borehole strain meters, tiltmeters, and satellite geodesy (continuous GPS and InSAR technologies in particular), it has become increasingly clear that stress changes resulting from sufficiently large earthquakes are capable of triggering a detectable response including earthquake activity, aseismic slip, and changes in volcanic and hydrothermal/hydrologic systems over a surprisingly wide range of distances and timescales. As currently understood, earthquake interactions are generally placed under one of the three rather general stress transfer modes: (1) static, (2) quasistatic, or (3) dynamic.

4.11.1.2 Static and Quasistatic Stress Triggering

Research on earthquake interactions associated with the change in the static stress field (static stress triggering) from just before an earthquake to shortly after the dynamic stresses from the seismic waves have decayed away began in the late 1960s and blossomed in the early 1990s with the growing availability of high-quality seismic and deformation data against which to test the patterns predicted by elastic dislocation models (Das and Scholz, 1981; Harris, 1998; **Chapter 4.10**, this volume; King and Cocco, 2001; Stein, 1999). Shortly thereafter, extensions of the dislocation models were developed to account for gradual stress changes

(quasistatic stress triggering) associated with viscous relaxation of the plastic lower crust and upper mantle in response to the sudden dislocation (an earthquake) across a fault in the overlying brittle crust (Pollitz and Sacks, 2002). Both the static and quasistress interaction modes appeal to the permanent stress change produced by the triggering earthquake or external loads including filling of reservoirs, ocean and glacier loading, (e.g., Harris, 1998; Hudnut et al., 1989; Kilb et al., 2002; Luttrell and Sandwell, 2010; Stein and Lisowski, 1983).

Because static stress changes resulting from an earthquake decay relatively rapidly with distance (as $\sim\Delta^{-3}$, where Δ is distance from the epicenter), their triggering potential is generally regarded as limited to one or two source dimensions from a given earthquake. Viscoelastic relaxation following a large crustal earthquake is largely confined to the lower crust and asthenosphere such that quasistatic stress changes propagate as a two-dimensional stress change. Quasistatic stress changes thus decay more slowly with distance (as $\sim\Delta^{-2}$), their triggering potential thus extends to greater distances than static stress changes, and the relatively low viscoelastic propagation speeds result in delayed triggering (Pollitz and Sacks, 2002). These two interaction modes are the focus of King (**Chapter 4.10**, this volume).

4.11.1.3 Dynamic Stress Triggering

The amplitudes of dynamic stresses propagating as seismic waves decrease relatively slowly with distance (as $\sim\Delta^{-2}$ for body waves and $\sim\Delta^{-3/2}$ for surface waves), and thus, their triggering potential extends from the near field (the aftershock zone) to much greater distances than either the static or the quasistatic stress changes as illustrated in **Figure 1** (Kilb et al., 2000). Amplification by radiation directivity, which is not a factor in either static

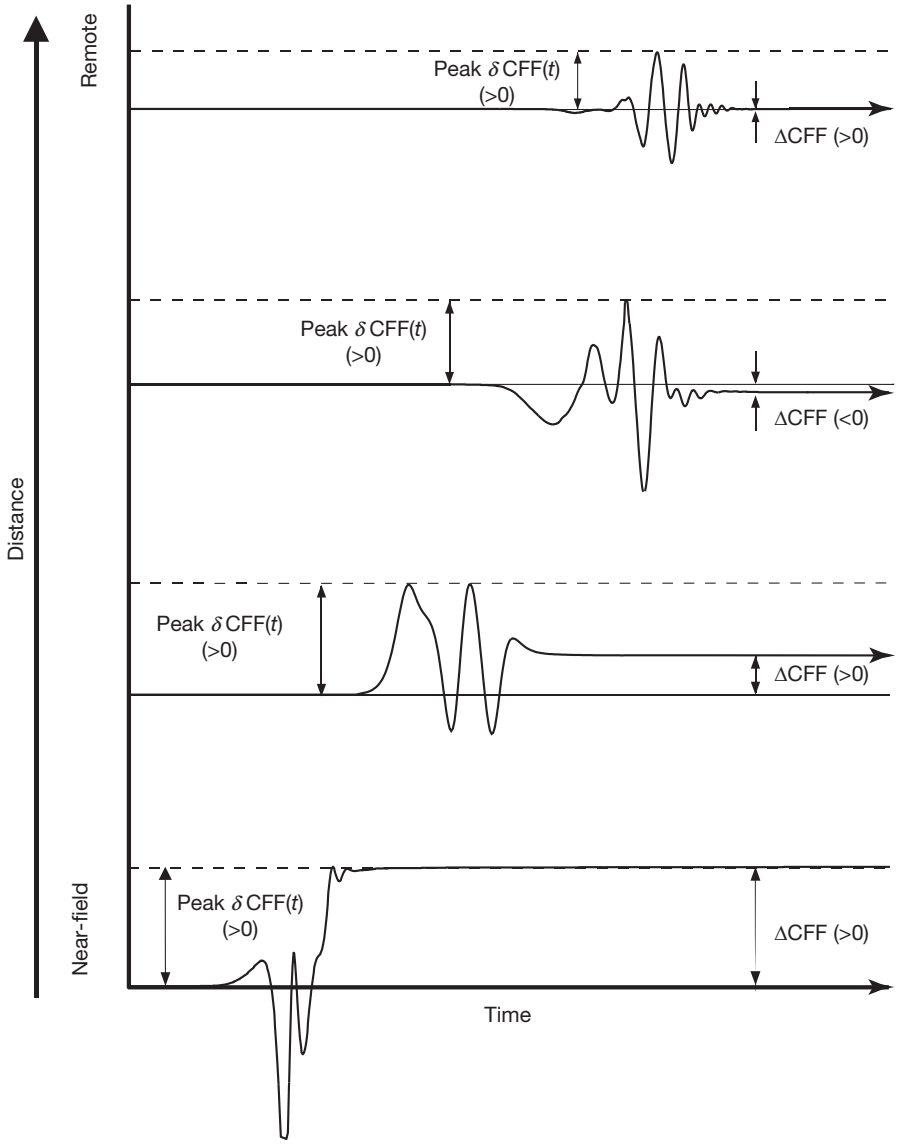


Figure 1 Diagram from Kilb et al. (2000) illustrating differences between static and dynamic stresses with distance in terms of idealized time histories for peak changes in dynamic (δCFF(t)) and static (ΔCFF) Coulomb stresses. δCFF(t) is the time-dependent version of ΔCFF (eqn [1]). The double arrows show peak values for δCFF(t) and ΔCFF. Adapted from Kilb D, Gomberg J, and Bodin P (2000) Triggering of earthquake aftershocks by dynamic stresses. *Nature* 408: 570–574, with permission.

or quasistatic stress changes, can further enhance the amplitudes of dynamic stresses in the case of mainshocks with asymmetric fault ruptures. Dynamic stresses are oscillatory, however, alternatively nudging the local stress field closer to and further from the local Coulomb failure stress over a range of frequencies leaving no permanent stress change. Indeed, it seems that Coulomb failure alone is not capable of explaining the full spectrum of observed dynamic triggering modes. In a number of cases, remotely triggered seismicity appears to be a secondary response to some more fundamental, aseismic process involving fluid activation or creep that was locally stimulated by the passing dynamic stresses. With periods of 12 and 24 h, the diurnal and semidiurnal tidal stresses induced in the solid Earth by the gravitational attraction of the sun and moon fall near the boundary between dynamic and quasistatic stresses (the inertial term is negligible), and they represent a lower bound on the frequency spectrum for dynamic triggering.

4.11.1.4 Triggered or Coincidental?

Conceptually, dynamic triggering is based on the inference of a causal link between dynamic stresses, such as seismic waves from a distant earthquake (event 'a'), and the onset of a local response as those stresses propagate through a site (event 'b'). Credible evidence in support of such a causal link emerges with multiple, well-documented observations that event 'b' follows event 'a' within some 'reasonable' time interval, Δt (**Figure 2**). Statistical methods, if carefully formulated, can add backbone to malleable qualifiers such as 'reasonable' and 'sufficiently,' thereby enhancing objective credibility that the temporal sequences 'a' and then 'b' cannot be dismissed as chance coincidences between random, statistically independent processes. Credible physical models that link calculable stress changes (whether static, quasistatic, or dynamic) to failure processes in rock volumes at the site of a putative triggered response will certainly add weight to statistical evidence for causal links.

A number of tests for the statistical significance of dynamic triggering of earthquake sequences appear in the literature, the most common of which is the β-statistic of Mathews and Reasenberg (1988) and the ZMAP software package, which incorporates the β-statistic (Wiemer, 2001). Statistical tests provide permissive but not sufficient evidence for a causal link between passing dynamic stresses and the onset of a local activity. Estimating the likelihood that an isolated event is triggered poses a special challenge in this regard as does a suspected delayed response as the delay time increases. Some of the most compelling examples of dynamic triggering involve an abrupt increase in seismicity rate at multiple sites as in the case of the Landers earthquake (**Figure 2**) or the modulation of bursts of earthquakes or tectonic tremor in phase with specific components of passing surface waves (e.g., Hill et al., 2013; Rubinstein et al., 2007; West et al., 2005). There is little question of a causal link in such cases, although questions may remain over the physical processes involved.

4.11.2 Early Inferences on Dynamic Triggering and Preinstrumental Examples

In an unpublished note dated 1955, Charles Richter wrote *The reverse effect – a major earthquake triggering a minor shock – is most probably [sic] within the immediate aftershock are [sic], but essentially by elastic wave propagation may set off action at a greater distance. If the distant effect is large enough, it may itself act as a trigger, so that there may be relay action, in which some of the later events are larger* (Hough, personal communication). Many of the early published inferences on dynamic triggering were associated with investigations into acoustic emissions and high-frequency seismic noise (e.g., Armstrong and Stierman, 1989; Nicolaev and Troitskii, 1987 and references therein, Galperin et al., 1990), and concerns over the possibility that large, underground nuclear explosions might trigger a damaging earthquake (Emiliani et al., 1969). At the time, these studies were greeted within the scientific community by attitudes ranging from mild interest to strong skepticism. Skepticism prevailed largely because of questions regarding the statistical significance of isolated observations together with the lack of a compelling physical model. The fact that some of these studies were a part of efforts focused on even more controversial topics such as anomalous animal behavior induced by acoustic emissions as a means of earthquake prediction (e.g., Armstrong and Stierman, 1989) probably contributed to the skeptical reception within the mainstream scientific community.

In a retrospective effort to find anecdotal evidence of preinstrumental dynamic triggering, Hough (2001) and Hough et al. (2003) preformed an extensive search through archive

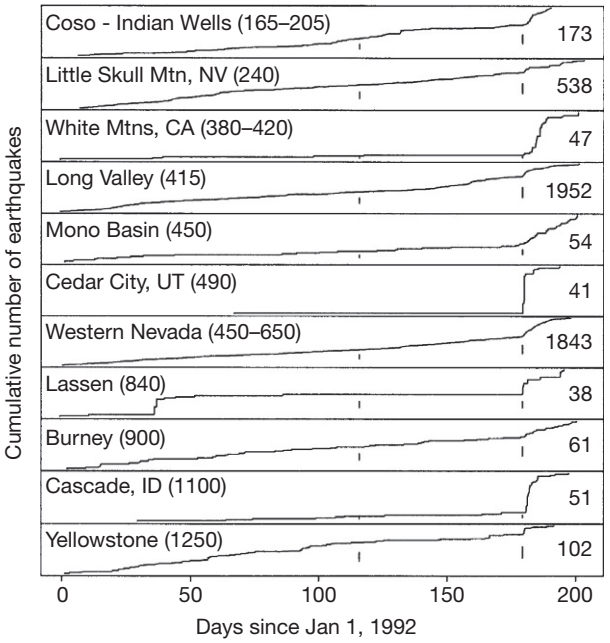


Figure 2 Evidence for remote dynamic triggering by seismic waves from the $M_w = 7.4$ Landers, CA, earthquake on 28 June 1992 in a plot of cumulative number of locatable earthquakes at 11 sites in the western United States for 211 days beginning with 1 January 1992. Numbers in parentheses are distances in kilometers from the Landers epicenter. Numbers at the right indicate the total number of earthquakes in the respective areas for the entire time interval. The short vertical bars at day 114 mark the $M_w = 7.1$ Petrolia earthquake on 25 April 1992 and those at day 180 mark the $M_w = 7.4$ Landers earthquake. Reproduced from Hill DP, Reasenberg PA, Michael A, et al. (1993) Seismicity remotely triggered by the magnitude 7.3 Landers, California, earthquake. *Science* 260: 1617–1623, with permission.

newspapers and correspondence for felt reports of local earthquakes that may have been triggered at remote distances by the four $M_w > 7$ New Madrid, Missouri, earthquakes of 1811–12 or the $M_w > 7$ Charleston, South Carolina, earthquake in 1886. They uncovered evidence that two of the New Madrid $M_w > 7$ events may have triggered locally felt earthquakes to distances of 1000 km or more.

Steeples and Steeples (1996) and Meltzner and Wald (2003) investigated the distribution of aftershocks and possible triggered earthquakes for the great ($M_w \sim 7.8$) San Francisco earthquake in 1906. They found evidence for nine earthquakes in the magnitude range $M \sim 4$ to ~ 6 within the first 48 h after the 1906 mainshock. The most distant, a $M \sim 4$ earthquake in western Arizona at $\Delta \sim 900$ km, occurred during the seismic waves from the 1906 mainshock and is a likely candidate for dynamic triggering.

Early investigations of dynamic triggering, however, are hampered by limited data. The situation began changing with the rapid expansion of continuously recorded, telemetered seismic networks around the globe through the 1980s and 1990s coupled with improvements in real-time processing and a growing number of broadband, high-dynamic-range digital installations. The latter set the stage for capturing a range of interesting seismic signals missed by earlier, short-period instrumentation.

4.11.3 Instrumental Evidence for Dynamic Triggering

On 28 June 1992, the $M_w = 7.3$ Landers earthquake ruptured a 70 km length of the Mojave Desert in southern California. Over the next few hours, it became clear that seismicity rates had increased at a number of sites across western North America at distances ranging from 200 to as much as 1250 km (17 source dimensions, see Figure 2). This large number of independent occurrences provided the first compelling evidence that remote triggering by dynamic stresses is indeed a viable process in the Earth (Gomberg, 1996; Hill et al., 1993). The Landers mainshock resulted from a unilateral rupture propagating to the north-northwest along a series of north-northwest striking dextral fault segments. All of the recognized sites of dynamic triggering were north of the Landers epicenter (Figure 3(b)), suggesting that amplification enhanced by rupture directivity may influence the distribution of dynamic triggering – a suggestion supported by subsequent instances of dynamic triggering (Gomberg et al., 2001).

Seven years later, the $M = 7.1$ Hector Mine earthquake on 16 October 1999 ruptured a 40 km length of the Mojave Desert along a series of faults located just 20 km east of the Landers rupture (Figure 3). In this case, fault rupture was bilateral but with the dominant rupture direction to the south-southeast. The most energetic triggered response to the dynamic stresses from the Hector Mine mainshock was in the Salton Trough south of the epicenter (Figure 3(a)), thus adding weight to the idea that rupture directivity and dynamic stress amplitudes are important factors in determining the distribution of dynamic triggering (Gomberg et al., 2001; Hough and Kanamori, 2002).

The $M_w = 7.9$ Denali Fault earthquake on 3 November 2002, which was centered 65 km east of Denali National Park, Alaska, resulted from a complex rupture with maximum

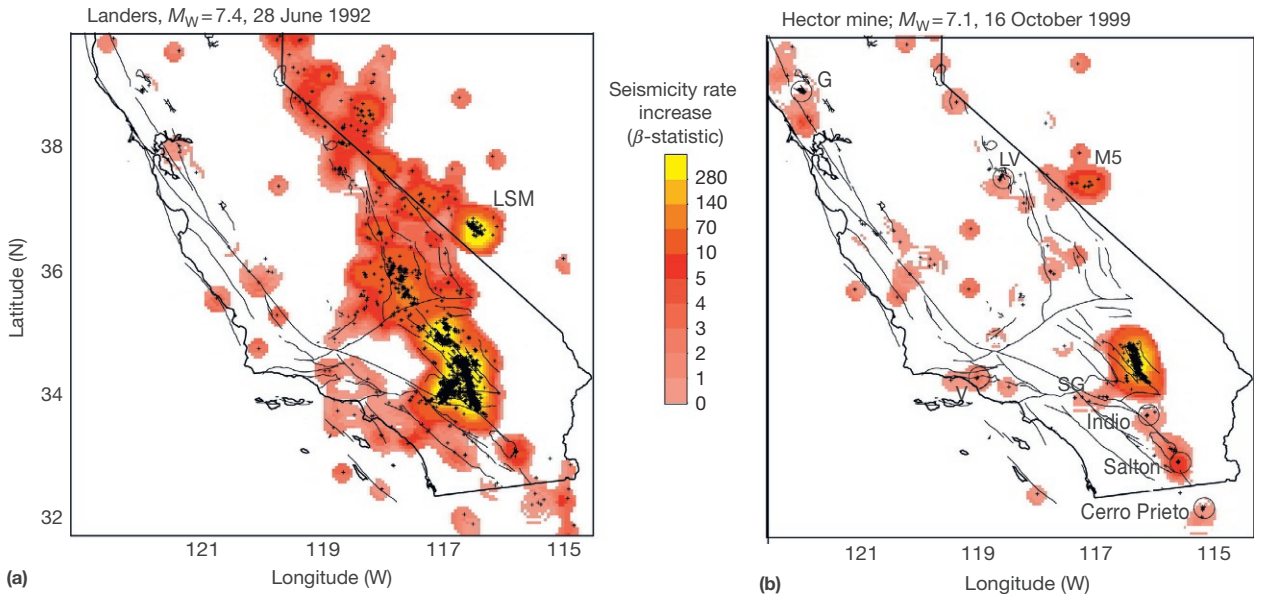


Figure 3 β-statistic maps illustrating seismicity rate increases associated with the $M_w = 7.4$ Landers earthquake on 28 June 1992 (a) and $M_w = 7.1$ Hector Mine earthquake on 16 October 1999 (b). The plots show a spatially smoothed β-statistic over dimensions of 20 km for $M \geq 2.0$ earthquakes from the Northern California Seismic Network (NCSN) and Southern California Seismic Network (SCSN) catalogs. The rate change was calculated for a 2-week period following the respective mainshocks with respect to the background period 1987 through mid-1992 (pre-Landers). Thin lines are major fault traces. Circles in (b) mark sites of activity triggered by the Hector Mine earthquake. M5 in (b) reflects aftershock activity to a $M = 5.3$ earthquake that occurred ~2.5 months prior to the Hector Mine earthquake. G, the Geysers geothermal area; LV, Long Valley Caldera; SG, San Gorgonio Pass; LSM, Little Skull Mountain. Modified from Gomberg J, Reasenberg PA, Bodin P, and Harris R (2001) Earthquakes triggering by seismic waves following the Landers and Hector Mine earthquakes. *Nature* 411: 462–465, with permission.

surface displacements reaching 8.8 m (Eberhart-Phillips et al., 2003; Rowe et al., 2004). The unilateral rupture produced pronounced directivity with peak amplification of dynamic stresses directed to the southeast through British Columbia and the western United States. With the exception of the Katmai volcanic cluster 740 km southwest of the epicenter, all recognized sites of dynamic triggering were located southeast of the epicenter in the sector of maximum directivity that spanned the areas responding to the Landers and Hector Mine earthquakes (Figure 4 in Hill and Prejean, 2007). In all cases, the onset of dynamic triggering developed as a seismicity rate increase during the passage of the Denali Love and Rayleigh waves (e.g., see **Figure 4**). In three cases, Mount Rainier, Long Valley Caldera, and Yellowstone, delayed earthquake swarms followed the initial surge in triggered seismicity by $\Delta t \sim 2.5$ h, 23.5 h, and 8 days, respectively (Husen et al., 2004a,b; Prejean et al., 2004). Pankow et al. (2004) documented both rapid-onset triggering and delayed triggering along the Wasatch Front in central Utah following the Denali earthquake.

The rapidly accumulating literature on dynamic triggering in the 20 years since the Landers earthquake makes it clear that dynamic triggering at remote distances is a much more widespread and a much richer phenomenon than initially envisioned. In the following three sections, we highlight some general aspects of the triggered response to dynamic stresses from large, distant earthquakes as currently understood through (1) brittle failure in the seismogenic crust, (2) fluid activation in the hydrologic domain and volcanic/geothermal systems, and (3) shear failure as the source of tectonic (non-volcanic) tremor in the transition from brittle–plastic behavior along major plate-boundary faults.

4.11.3.1 Widespread Dynamic Triggering in the Seismogenic Crust

Velasco et al. (2008) examined the records from 500 globally distributed broadband seismic stations for evidence of remote dynamic triggering by 15 $M > 7$ earthquakes occurring over the 18-year span from 1990 to 2007. They found that, out of the

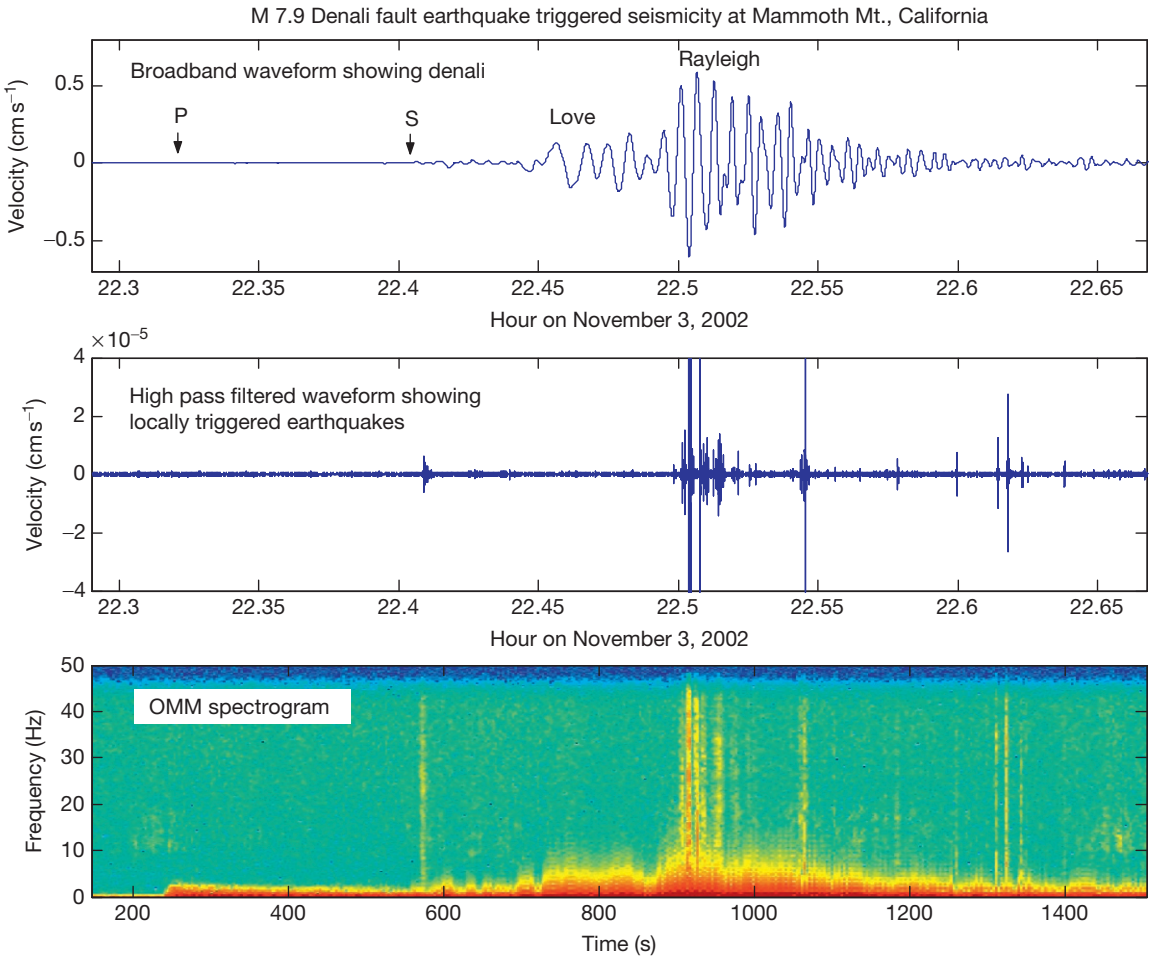


Figure 4 Examples of triggered seismicity detected in the surface-wave coda from the Denali Fault earthquake at Mammoth Mountain. Top trace is the broadband seismogram from station OMM located 4 km SE of Mammoth Mountain. Middle and bottom panels show the high-pass-filtered version of the OMM seismogram and a spectrogram of the unfiltered seismogram. Modified from Prejean SG, Hill DP, Brodsky EE, et al. (2004) Remotely triggered seismicity on the United States west coast following the Mw = 7.9 Denali Fault earthquake. *Bulletin of the Seismological Society of America* 94: S348–S359, courtesy of the Seismological Society of America.

15 $M>7$ mainshocks studied, 12 were associated with significant increases in the detection of smaller earthquakes in the vicinity of the respective seismic stations all located at distances $>10^{\circ}$ ($\sim$1110 km) from the mainshock epicenters. Their analysis revealed that, on average, the local earthquake rate increased above background by 38% with the passage of the Love wave and 60% with the passage of the Rayleigh wave and that remote triggering occurred in all tectonic environments. Parsons and Velasco (2011) found no evidence for $M>5$ earthquakes triggered during the passage of surface waves at distances beyond two to three rupture lengths from any of the $M>7$ earthquakes having occurred in the previous 30 years. In a paper published just a year later, Pollitz et al. (2012) documented a fivefold increase in the global rate of $M>5.5$ earthquakes at distances <1500 km from the triggering mainshock during the first 6 days following the $M_w=8.6$ Indian Ocean earthquake on 11 April 2012. Only one of these $M>5.5$ events (a $M=5.5$ earthquake beneath the western Aleutian Islands) occurred during the passing seismic waves, however; the remainder were delayed by hours to several days. This underscores the rapid evolution of our perspective on dynamic triggering with time as new data become available (**Tables 1** and **2**; **Figure 5**).

Gonzalez-Huizar et al. (2012) found evidence that the 2011 $M_w=9.0$ Tohoku-Oki earthquake triggered crustal seismicity in South Dakota (the United States), Russia, China, Ecuador, and Mexico. The latter involved the abrupt onset of an earthquake swarm in a local spreading center in the Gulf of California that included a pair of $M=5.2$ earthquakes in same area as the $M_w=7.0$ earthquake triggered by the 2012 $M_w=$Indian Ocean earthquake (Pollitz et al., 2012). Other examples from around the globe include the Netherlands (Camelbeeck et al., 1994), the Yammouneh fault along the Syria–Lebanon border (Mohamad et al., 2000), continental Greece (Brodsky et al., 2000), the Valley of Mexico (Singh et al., 1998), Taiwan (Wen et al., 1996), east-central China (Jiang et al., 2010; Wu et al., 2011), Japan (Miyazawa, 2011, 2012), geothermal areas in California (Brodsky, 2006; Brodsky and Prejean, 2005; Peng et al., 2011a,b), and Queensland, Australia (Gonzalez-Huizar and Velasco, 2011). The following examples are particularly noteworthy.

The epicenters of $M_w=6.5$ and 6.4 earthquakes in the South Iceland Seismic Zone on 17 and 21 June 2000, respectively, were separated by just 17 km in an east–west direction. The first and largest of these earthquakes triggered widespread seismicity on the Reykjanes Peninsula to distances of $\sim$100 km to the west and significant slip on at least three faults in the Reykjanes Peninsula, while the second, slightly smaller earthquake, although 17 km further west, triggered little response (Arnadottir et al., 2004; Daniel et al., 2008). The triggered seismicity included three $M\sim5$ earthquakes that followed the $M_w=6.5$ event by seconds to minutes. Arnadottir et al. (2004) concluded that these $M\sim5$ earthquakes were beyond the range for Coulomb failure due to static stress changes from the

Table 1 Source earthquakes for cited instances of dynamic triggering

Earthquake	*Magnitude*	*Date*	*Depth (km)*	*Slip*	*Representative references*
New Madrid, MO, USA ($4M>7.5$)	M 7.5–7.7	1811–12	<15	SS	Hough (2001)
Charleston, NC, USA	M 7.3	01 September 1886	<15	SS	Hough et al. (2003)
Kangra, India	M_s 7.8	04 April 1905	<20	T	Hough et al. (2005)
San Francisco, CA, USA	M 7.8	18 April 1906	<15	SS	Meltzner and Wald (2003)
Roer Valley, the Netherlands	M_w 5.4	13 April 1992	<15	SS	Camelbeeck et al. (1994)
Landers, CA, USA	M_w 7.4	28 June 1992	<15	SS	Hill et al. (1993)
Gulf of Aqaba	M_w 7.3	22 November 1995	<15	SS	Mohamad et al. (2000)
Hector Mine, CA, USA	M_w 7.1	16 November 1999	<15	SS	Gomberg et al. (2001)
Izmit, Turkey	M_w 7.6	17 August 1999	15	SS	Brodsky et al. (2000)
Chi-Chi, Taiwan	M_w 7.7	20 September 1999	5	T	Miyazawa et al. (2005)
Iceland	M_w 6.5	17 June 2000	6	SS	Arnadottir et al. (2004)
Western Tottori, Japan	M_w 6.7	06 November 2000	15	SS	Di Carli et al. (2008)
Bhuj, western India	M_w 7.6	26 January 2001	10	T	Surve and Mohan (2012)
Kunlun, China	M_w 7.8	14 November 2001	10	SS	Wu et al. (2011)
Denali Fault, AK	M_w 7.9	11 March 2002	5	SS	Prejean et al. (2004)
Palermo, Italy	M_w 6.0	09 June 2002	10	T	Walter et al. (2009)
Tonga	M_w 7.6	08 September 2002	665	T	Tibi et al. (2003)
Tokachi–Oki, Japan	M_w 8.3	25 September 2003	27	T	Miyazawa and Mori (2005)
SW Siberia, Russia	M_w 7.3	27 September 2003	17	SS	Miyazawa and Mori (2005)
Sumatra, Indonesia	M_w 9.1	26 December 2004	30	T	Lay et al. (2005)
Southern Greece	M_w 6.8	09 June 2006	66	T	Cannata et al. (2010)
Yogyakarta, Indonesia	M_w 6.3	26 June 2006	10	SS	Mori and Kano (2009)
Solomon Islands	M_w 8.1	01 April 2007	10	T	Miyazawa et al. (2008)
Wenchuan, China	M_w 7.9	12 May 2008	19	T	Miyazawa et al. (2008)
Halmahera, Indonesia	M_w 6.6	11 September 2008	98	T	Lin (2012)
Kepulauan Talaud, Indonesia	M_w 7.2	11 February 2009	20	T	Jousset and Rohmer (2012)
Maule, Chile	M_w 8.8	27 February 2010	35	T	Zigone et al. (2012)
Tohoku–Oki, Japan	M_w 9.0	11 March 2011	29	T	Gonzalez-Huizar et al. (2012)
Indian Ocean	M_w 8.6	11 April 2012	20	SS	Pollitz et al. (2012)

Slip: SS, strike-slip; T, thrust or reverse.

Table 2 Reported instances of remote dynamic triggering in the brittle domain

Site	*Triggered* M_{max}	*Regime*	*Mainshock*		*References*
			M_{max}	*Distance (km)*	
North America					
Mt. Wrangell, AK	<1	V	9.0	~11 000	28
Katmai, AK	2.3	G, V	7.9	115–740	21, 43
South BC, Canada	N/A	C	7.9	1800–2200	9
Mt. Rainier, WA	<1	V	7.9	3108	23
Geysers, CA	<3	E, G, V	6.5–7.9	202–3120	7, 23, 17
Coso, CA	3.2	E, G, V	< 7.9	165–3660	23
Long Valley, CA	3.4	E, G, V	7.4–7.9	414–3454	8, 23
Mammoth Mountain, CA	$M<2$	E, G, V	7.2–7.9	420–3454	2, 16
Lassen Peak, CA	2.8	E, V	7.4	840	10
Burney, CA	2.8	E	7.4	900	10
Salton Sea area, CA	4.7	E, V, G	7.1	120–150	14
Central and south CA	5?	E, C	5.8–6.1	70–120	13
California	–	E, C	2.1–5.5	>3000	41
Offshore S. CA	2.5	E	7.9	4003	23
W. Nevada	~4	E, G	7.4	450–650	1
W. Arizona (1906)	~4	E	7.8	900	29, 30
Little Skull Mountain, NV	5.6	E	7.4	240	1
Yellowstone, WY	3.0	E, G, V	7.4–7.9	1250–3100	15
South Dakota	<3	I	9.0	8603	33
Wasatch Front, UT	3.2	E, G	7.4–7.9	3000–3500	22
Cascade, ID	1.7	E, G	7.4	1100	9
The eastern United States (1811–12)	~5	C	$M>7$	~1000	12
South and Central America					
Cerro Prieto, Mexico	4.1	E, V, G	7.1	260	6
Valley of Mexico	~4	E, G, V	7.6–8.0	303–588	24
Gulf of CA, Mexico	5.2	E	9.0	~11 000	33
Ecuador	–	E/C	9.0	~14 000	33
Asia					
Aso, Japan	~2	E, V	7.1–7.7	900–2213	18
Japan	≤5	V, C	9.0	1350	37, 39
Hakone caldera, Japan	4.8	V	9.0	450	40
Iwo Jima, Japan	<2	IA, G, V	7.1–8.0	1228–2002	26
Taiwan region	>4	?	6.5–7.1	138–2959	27
Nankai Trough, Japan	N/A	S	7.3–8.1	900–4000	19
Tonga Trench	5.9–7.7	S	7.1–7.6	260–290	25
Central China	–	E	7.8–9.0	1497–19 140	33, 34, 35, 42
Queensland, Australia	–	C	7.5–8.4	15 000	36
Europe					
SISZ, Iceland	~5	E, G	6.5	80–100	2
Roer Valley, Holland	3.7	E	5.4	40	4
Greece	>3.5	E	7.4	400–1000	3
Turkey	>2.8–3.0	E	7.6	20–80	5
Paris Basin, France	<1		7.2	12 000	38
Syria–Lebanon border	3.7	C	7.3	500	20
Russia	–	C	9.0	2971	33
Global studies					
Global waveforms	<1	–	>7	>10°	31
Global catalogs	>5.5	–	8.6		32

Tectonic regimes: E, extensional, transtensional; C, convergent, transpressional; I, intraplate; G, geothermal; V, volcanic; IA, island arc; S, subcrustal subduction zone; SISZ, South Iceland Seismic Zone.

References: 1. Anderson et al. (1994); 2. Arnadottir et al. (2004); 3. Brodsky et al. (2000); 4. Camelbeeck et al. (1994); 5. Durand et al. (2013); 6. Glowacka et al. (2002); 7. Gomberg (1996); 8. Gomberg et al. (2001); 9. Gomberg et al. (2004); 10. Hill et al. (1993); 11. Hill et al. (1995); 12. Hough (2001); 13. Hough (2005); 14. Hough and Kanamori (2002); 15. Husen et al. (2004a,b); 16. Husker and Brodsky (2004); 17. Johnston et al. (2004b); 18. Miyazawa et al. (2005); 19. Miyazawa and Mori (2005); 20. Mohamad et al. (2000); 21. Moran et al. (2004); 22. Pankow et al. (2004); 23. Prejean et al. (2004); 24. Singh et al. (1998); 25. Tibi et al. (2003); 26. Ukawa et al. (2002); 27. Wen et al. (1996); 28. West et al. (2005); 29. Steeples and Steeples (1996); 30. Meltzner and Wald (2003); 31. Velasco et al. (2008); 32. Pollitz et al. (2012); 33. Gonzalez-Huizar et al. (2012); 34. Jiang et al. (2010); 35. Wu et al. (2011); 36. Gonzales-Huizar and Velasco (2011); 37. Miyazawa (2011); 38. Jousset and Rohmer (2012); 39. Hirose et al. (2011); 40. Yukutake et al. (2011); 41. Van der Elst and Brodsky (2010); 42. Wu et al. (2012). Power et al. (2001).

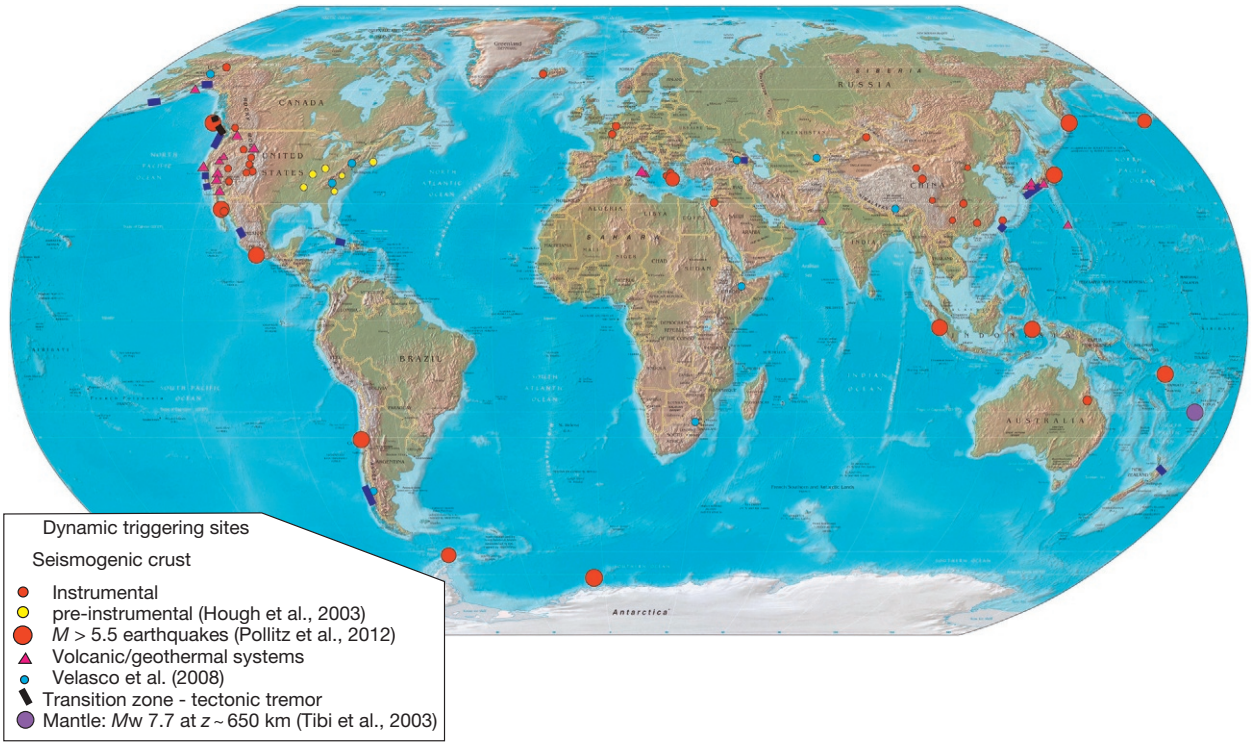


Figure 5 Map indicating the distribution of published instances of dynamically triggered earthquakes referred to in this chapter. Triggered seismicity in the brittle crust indicated by red circles for instrumentally recorded sites and orange circles for inferred, preinstrumental sites (Hough et al., 2003). Red triangles indicate volcanic/geothermal sites showing a triggered response. Large red circles indicate the $M > 5.5$ events that occurred in the first 6 days after the 2012 $M_w = 8.6$ Indian Ocean earthquake (Pollitz et al., 2012). Black bars indicate documented sites of triggered tectonic tremor. Blue circles indicate instrument clusters that showed an increase in seismicity rate following 12 out of 15 $M > 7$ earthquakes between 1990 and 2008 (Velasco et al., 2008). Large purple circle indicates the 650 km deep $M_w = 7.7$ earthquake triggered by the 202 $M_w = 7.6$ earthquake beneath the Tonga subduction zone (Tibi et al., 2003).

$M_w = 6.5$ mainshock and thus that they were likely triggered by dynamic stresses. Their analysis of continuous GPS geodetic data spanning the Reykjanes Peninsula indicates that the geodetic moment for the second of the triggered $M \sim 5$ earthquakes ($\Delta \sim 78$ km) was an order of magnitude larger than its seismic moment and that the Coulomb stress change from this largely aseismic, slow earthquake was sufficient to trigger the third and most distant $M \sim 5$ event some 4 min later through quasistatic stress transfer.

A compelling example in support of dynamic triggering being confined to crustal volumes in a near-critical stress state comes from data recorded during a salt cavern collapse at the eastern edge of the Paris Basin, Lorraine, France (Jousset and Rohmer, 2012). The cavern was created by a salt-mining process that involves injecting freshwater into sedimentary salt layer through boreholes in one area and extracting brine from another some distance away. The cavern grew as the salt was extracted, with the process monitored by tracking ground deformation and seismicity. In this case, the mine operator decided to proceed to controlled cavern collapse. Jousset and Rohmer (2012) examined the local earthquake catalog for signs of dynamic triggering by $M > 7$ earthquakes in the year before the planned collapse and found none. As the dimensions of the cavern approached a critical state, however, seismic waves from the 2009 $M_w = 7.2$ Kepulauan Talaud earthquake (Indonesia) triggered the onset of a seismicity sequence that culminated in cavern collapse 3 days later. In their analysis of the sequence, Jousset and Rohmer (2012) concluded that Rayleigh wave dynamic stresses of 1–2 kPa from the $M_w = 7.2$ earthquake were the dominant triggering agent.

Not all instances of remotely triggered brittle failure are confined to the brittle crust. Peng and Walter (2013) found that the 2010 $M_w = 8.8$ Maule, Chile, earthquake triggered ice quakes in the glacier near the Howard Nunataks and Ellsworth Mountains, Antarctica. Based on the nearby broadband station HOWD, they concluded that the triggered activity coincided with dilatational stresses associated with both *P*-waves and Rayleigh waves. At the other extreme, a $M_w = 7.6$ earthquake on 19 August 2002 at a depth of ~665 km beneath the Tonga subduction zone was followed by M_b 5.9 and $M_w = 7.7$ earthquakes in the same depth range with delays of 2.2 and 7.4 min, respectively (Tibi et al., 2003). The latter two earthquakes occurred in a previously aseismic volume located ~290 km southwest of the initial $M_w = 7.6$ event. Tibi et al. (2003) found that the static Coulomb stress change at this distance from the initial $M_w = 7.6$ earthquake is small and concluded that these earthquakes were most likely dynamically triggered by the body waves from the $M_w = 7.6$ earthquake.

4.11.3.2 Dynamic Triggering of Volcanic/Geothermal and Hydrologic Systems

Recognition that volcanic and geothermal areas appear to be particularly susceptible to dynamic triggering of local earthquake activity by seismic waves from distant earthquakes heightened interest in the broader topic of earthquake–volcano interactions (Delle-Donne et al., 2010; Hill et al., 2002; Linde and Sacks, 1998; Manga and Brodsky, 2006). It has long been recognized that shallow hydrologic systems respond to seismic waves from large earthquakes (**Chapter 4.12**, this volume; Manga and Wang, 2007; Richter, 1958; Vohis, 1967; Wang and Manga, 2009). In the case of earthquake–volcano interactions, static or quasistatic stress transfer can work in both directions with the stress change associated with a magmatic intrusion or eruption triggering slip on a nearby fault or vice versa (e.g., Chesley et al., 2012; Nostro et al., 1998; Walter, 2007). Our focus here, however, remains on dynamic triggering.

Many of the areas responding to the 1992 $M_w = 7.3$ Landers earthquake, the 1999 $M_w = 7.1$ Hector Mine, and the 2002 $M_w = 7.1$ Denali Fault earthquakes were volcanic or geothermal including Long Valley and Yellowstone Calderas. The response of Yellowstone Caldera to the Denali Fault earthquake in particular included pronounced transient changes in the hydrothermal and geyser/hot-spring systems (Husen et al., 2004b; Pankow et al., 2004). The response of Long Valley Caldera to the Landers, Hector Mine, and Denali Fault earthquakes was dominated by a largely aseismic deformation transient with the deformation moment significantly greater than the cumulative seismic moment of the triggered seismicity (**Figure 6**).

The Geysers geothermal field in the northern California Coast Ranges seems particularly susceptible to dynamic triggering. It responded with increased seismicity rates to dynamic stresses from eight regional earthquakes with magnitudes $M = 6.6–7.7$ at distances $\Delta = 212–2500$ km between 1988 and 1994 (Gomberg and Davis, 1996; Stark and Davis, 1995) and the 1999 Hector Mine earthquake (Gomberg et al., 2001), and the 2005 $M_w = 7.2$ Mendocino earthquake (Brodsky, 2006) and to the teleseismic waves from the 2002 Denali Fault earthquake located 3120 km to the northwest (Prejean et al., 2004).

Surve and Mohan (2012) documented increased earthquake swarm activity in the Western Ghats geothermal area in the Deccan volcanic province of India following the 2001 $M_w = 7.6$ Bhuj earthquake, which was located in western India 500 km to the northwest. The onset of swarm activity began ~2.5 h after the Bhuj earthquake and continued with three bursts in activity over the next 2 months. In contrast, Sanchez and McNutt (2004) see intermediate-term declines in seismicity beneath Mt Warngell and Mt Veniaminof volcanoes in Alaska following the 2002 $M_w = 7.9$ Denali Fault earthquake.

Harrington and Brodsky (2006) found that, except for volcanic systems on Kyushu, there was surprisingly little evidence for dynamic triggering in Japan by distant earthquakes. Miyazawa et al. (2005) described evidence for multiple increases in isolated tremor and microearthquakes at Aso volcano in central Kyushu following large, distant earthquakes including the $M_w = 7.7$ Chi-Chi, Taiwan, earthquake on 20 September 1999. In 2011, however, the 11 March $M_w = 9.0$ Tohoku-Oki earthquake triggered widespread seismicity across Japan, much of which was located in volcanic or geothermal areas (Hirose et al., 2011; Miyazawa, 2011). A seismicity increase in Hakone caldera at an epicentral distance of 450 km from the Tohoku-Oki epicenter coincided with the passing surface waves with peak dynamic stresses of 0.6 MPa – an order of magnitude larger than the estimated ~0.04 MPa static stress change at the caldera associated with the Tohoku-Oki mainshock (Yukutake et al., 2011, 2013).

The island of Iwo Jima in the Volcano Islands group situated between the Marianas and Izu island arcs some 1250 km south of Tokyo represents the resurgent dome of a large submarine caldera. Ukawa et al. (2002) examined Iwo Jima seismograms for 21 $M > 7$ earthquakes located within 3000 km of the island and found four instances when local seismicity increased during the passage of the surface waves from the distant earthquakes. Iwo Jima seismograms showed a similar triggered response to the surface waves from the great $M_w = 9.1$ Sumatra–Andaman Islands earthquake on 26 December 2004 and those from the 2011 $M_w = 9.0$ Tohoku-Oki earthquake (Ukawa, personal communication). In contrast, Okubo and Wolfe (2008) cited evidence for a significant decrease in the rate of 30 km deep long-period (LP) earthquakes beneath Mauna Loa volcano, Hawaii, a quiescence apparently triggered by dynamic stresses from the 2004 $M_w = 9.1$ Sumatra earthquake.

Volcanic unrest at Mount Etna on Sicily responded to dynamic stresses from the 8 January 2006 $M_w = 6.8$ earthquake in southern Greece and the 12 May 2008 $M_w = 7.9$ earthquake in eastern Sichuan, China (Cannata et al., 2010). The latter was followed by an abrupt decrease in the frequency content of ongoing seismic activity within the volcano and the onset of an explosive eruption 27 h later. An unusual cluster of three eruptions followed a $M_w = 5.9$ earthquake on 6 September 2002 centered ~50 km north of Palermo. The largest Mount Etna eruption in 55 years occurred on 28 October 2002 followed by a submarine degassing event near Panarea Island on 3 November 2002 and the largest effusive eruption of Stromboli in 17 years on 28 December 2002 all located within 150 km of the Palermo earthquake (Walter et al., 2009). The 2–4-month delay between the earthquake and the eruptions, however, leaves a causal link open to question.

The value of a well-monitored volcano with multiparameter instrumentation is underscored by the eruption of Shinmoedake volcano in Japan, which followed the 11 March 2011 $M_w = 9.0$ Tohoku-Oki earthquake by just 2 days. In this case, Yamazaki et al. (2011) concluded that the eruption was probably not triggered based on data from a nearby borehole strain meter showing no evidence for aseismic deformation nor a pressure increase in the magma chamber. Similarly, Mori and Kano (2009) concluded that evidence was at best ambiguous for a causal link between the $M_w = 6.3$ Yogyakarta earthquake at a distance of 250 km from the damaging and widely publicized Lusi mud volcano eruption that began 3 days later at a drill site in Sidoarjo, Indonesia. Indeed, Davies et al. (2010) provided compelling evidence that the Lusi eruption was the result of drilling operations at the drill site.

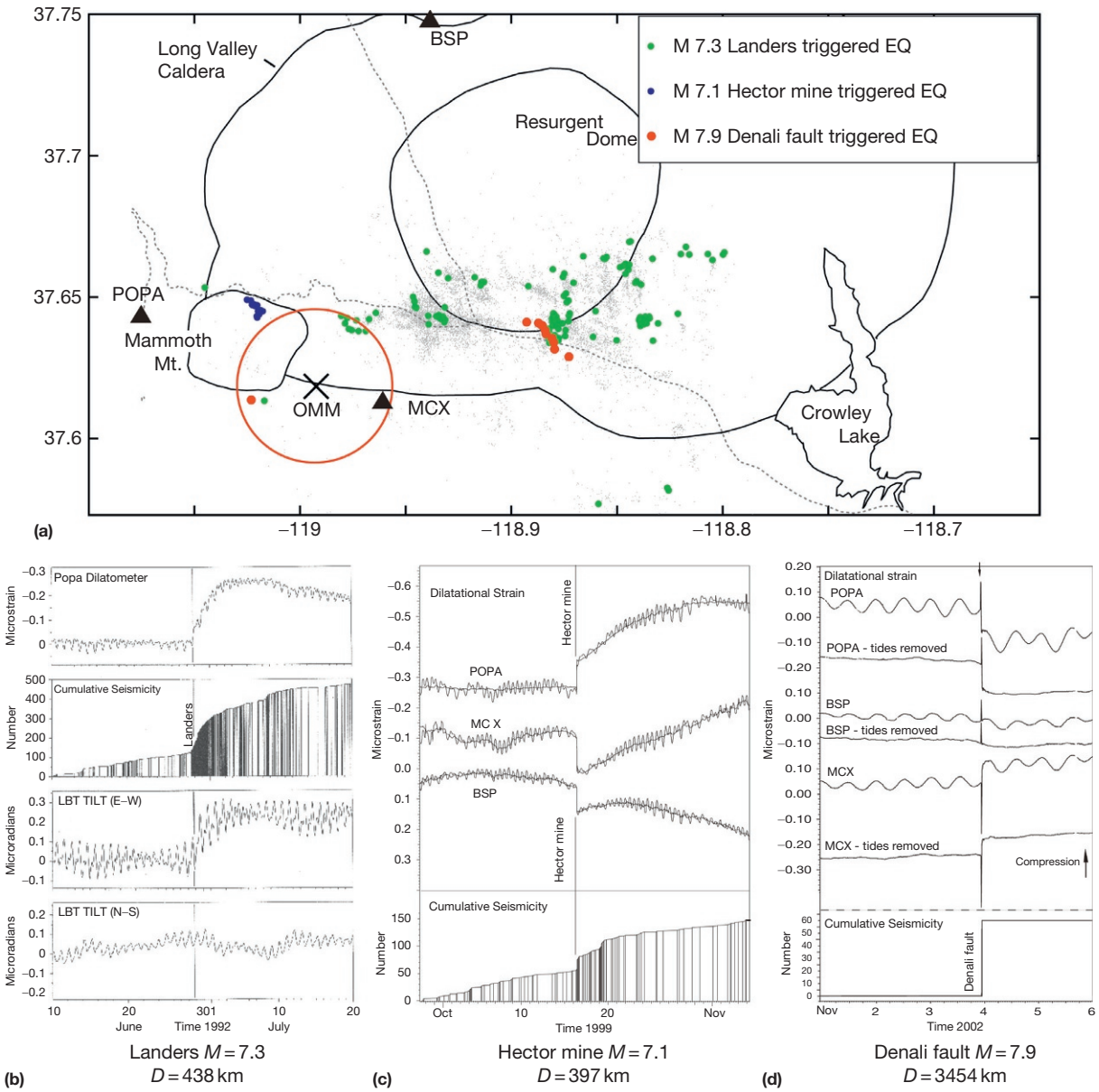


Figure 6 (a) Map of triggered seismicity beneath Long Valley Caldera and Mammoth Mountain for the Landers (green), Hector Mine (blue), and Denali Fault earthquakes (red). Gray dots show seismicity for 1997–98. The red circle centered on station OMM indicates area within which the Denali Fault-triggered earthquakes must be located based on S–P times. The single red dot is the epicenter for those earthquakes large enough for a multistation location. Modified from Prejean SG, Hill DP, Brodsky EE, et al. (2004) Remotely triggered seismicity on the United States west coast following the Mw 7.9 Denali Fault earthquake. *Bulletin of the Seismological Society of America* 94: S348–S359, courtesy of the Seismological Society of America. (b) Deformation transient triggered by the Landers earthquake as recorded on the POPA borehole dilatometer (top) and the long base tiltmeter (LBT). The bottom two panels show E–W and N–S tilt components, respectively. Cumulative number of earthquakes plotted in second panel. Adapted from Johnston MJS, Hill DP, Linde AT, Langbein J, and Bilham R (1995) Transient deformation during triggered seismicity from the 28 June 1992 Mw = 7.3 Landers earthquake at long valley volcanic caldera, California. *Bulletin of the Seismological Society of America* 85: 787–795, courtesy of the Seismological Society of America. (c) Deformation transient triggered by the Hector Mine earthquake as recorded on borehole dilatometers POPA, MCX, and BSP. Cumulative number of earthquakes plotted in bottom panel. Reproduced from Johnston MJS, Hill DP, and Pitt AM (2000) Strain transient recorded in the Long Valley Caldera during triggered seismicity from the October 16, 1999 M7.1 Hector Mine, California, earthquake. *EOS, Transactions of the American Geophysical Union* 81: WP1384. (d) Deformation transient triggered by the Denali Fault earthquake as recorded on borehole dilatometers POPA, MCX, and BSP. Strain records for each dilatometer are shown with solid Earth tides (top) and with tides filtered out (bottom). Cumulative number of earthquakes plotted in bottom panel. Adapted from Johnston MJS, Prejean SG, and Hill DP (2004a) Triggered deformation and seismic activity under Mammoth Mountain in long valley caldera by the 3 November 2002 Mw 7.9 Denali fault earthquake. *Bulletin of the Seismological Society of America* 94: S360–S369, courtesy of Seismological Society of America.

4.11.3.3 Dynamic Triggering of Tectonic (Nonvolcanic) Tremor in the Transitional Domain

Following the discovery of tectonic (nonvolcanic) tremor in Japan (Obara, 2002) and recognition of its relation to episodic slow slip on the Nankai and Cascadia subduction zone faults (Julian, 2002; Rogers and Draggert, 2003), the combination of episodic tremor and slip (ETS) in the transitional domain downdip of the seismogenic section of major plate-boundary faults is now recognized as an important component in the earthquake cycle for major, plate-boundary earthquakes (Beroza and Ide, 2011; Rubinstein et al., 2010; **Chapter 4.17**). Using cross correlation methods to obtain precise hypocentral locations of low-frequency earthquakes (LFEs) and recognizing their close association with slow slip events in Japan, Shelly et al. (2006) proposed that tectonic tremor results from sequences of LFEs with dominant energy in the 1–10 Hz frequency band, each of which represents an increment of shear failure along the megathrust fault plane at depths of 20–30 km.

Obara (2002) first recognized that levels of tremor activity beneath Japan can be modulated by dynamic stresses associated with seismic waves from both local and teleseismic earthquakes. Miyazawa and Mori (2005) applied a β-statistic analysis to tremor in the Nankai subduction zone and found a rate increase coincident with the Rayleigh waves from the $M_w = 8.1$ Tokachi–Oki earthquake on 25 September 2003, some 1000 km to the northwest. They found a similar response from a $M_w = 7.3$ earthquake in Siberia 40 h later (27 September). Subsequent studies by Miyazawa and Mori (2006), Miyazawa and Brodsky (2008), and Miyazawa et al. (2008) found triggered tremor in Nankai coincident with surface waves from the 2004 $M_w = 9.1$ Sumatra–Andaman Islands earthquake, the 2007 $M_w = 8.1$ Solomon Islands earthquake, and the 2008 $M_w = 7.9$ Wenchuan earthquake. As documented by Rubinstein et al. (2007), Rubinstein et al. (2009), Gomberg (2010), and Chao et al. (2013), surface waves from $M_w > 7$ earthquakes at teleseismic distances commonly trigger tremor on the Cascadia subduction zone off coastal Washington and Oregon. Tremor and ETS events on both subduction zones occur on low-angle megathrust fault planes dipping at 10°–15° at depths of 30–40 km.

Triggered tremor has subsequently been identified on a growing number of subduction zone megathrust fault planes with low-angle dips (10°–15°) at depths of 30–50 km. This includes subduction zones off the west coasts of Mexico, southern Chile, and the east coast of New Zealand (Chao et al., 2013; Fry et al., 2011Gonzalez-Huizar et al., 2012; Peng et al., 2013; Rubinstein et al., 2010; Zigone et al., 2012). Fry et al. (2011) identified tectonic tremor on the subduction zone beneath the east (Hikurangi) coast of the North Island, New Zealand. Ambient tectonic tremor and instances of triggered tremor have also been identified on several plate-boundary transform faults including the Parkfield section of the San Andreas Fault in central California (Guillhem et al., 2010; Hill et al., 2013; Peng et al., 2009, 2010a,b; Shelly et al., 2011) and some of its branches in northern and southern California (Chao et al., 2012; Gomberg et al., 2008; Wang et al., 2013; Yang and Peng, 2013); the Queen Charlotte Fault on Queen Charlotte Island, Canada; the Oriente fault in southern Cuba; and the Enriquillo–Plantain Garden fault in Haiti (Aiken et al., 2013a; Gonzales-Huizar et al., 2012; Peng et al., 2013). Triggered tremor has also been identified beneath the Central Range in Taiwan, although it is still not clear whether tremor occurred on a low-angle detachment fault at depths of 15–25 km (Chao et al., 2011; Peng and Chao, 2008) or high-angle thrust faults (Tang et al., 2010, 2013). Wech et al. (2012) identified tectonic tremor at depths of 25–45 km along the deep extension of the transpressional Alpine Fault along the west coast of the South Island of New Zealand, but as of yet anyway, it has no instances of triggered tremor (**Table 3**).

In most cases, triggered tremor on major subduction zone and transform faults begins with the arrival of the fundamental Love wave with periods ranging from 20 to 200+s and often continues well into the Rayleigh wave coda, in many cases with tremor amplitudes apparently modulated by the Rayleigh wave. Triggered tremor generally occurs in the same source volume that produces background tremor and ETS events, but it is usually somewhat more energetic than the background tremor (Peng et al., 2008; Rubinstein et al., 2007). The Nankai subduction zone represents an exception with most tremor activity commencing with the arrival of the Rayleigh wave. For surface waves arriving with near strike-parallel incidence on the megathrust fault plane (e.g., waves from the 2004 $M_w = 9.1$ Sumatra and the 2011 $M_w = 9.0$ Tohoku-Oki earthquakes on both Cascadia and Nankai, the 2003 $M_w = 8.1$ Tokachi–Oki earthquake on Nankai, and the $M_w = 7.9$ 2002 Denali Fault earthquake on Cascadia), Love waves should have a much higher potential for triggering reverse slip than Rayleigh waves. This is consistent with the Cascadia observations (Hill, 2012a) and the inference by Rubinstein et al. (2007) that Cascadia tremor results from small increments of reverse slip on the fault plane. Miyazawa and Brodsky (2008) and Miyazawa et al. (2008) suggested that in Nankai, the Rayleigh wave dilatational stress may have an important role in modulating tectonic tremor beneath southwestern Japan by modulating fracture-zone permeability and pumping fluids through vertical fractures in combination with a relatively high friction coefficient along the megathrust interface (also see Hill, 2012a). More recently, however, Enescu et al. (2012) reported clear evidence of Love wave triggered tremor in the Nankai subduction zone following the 2001 $M_w = 7.9$ Kunlun earthquake and the 2012 $M_w = 8.6$ Indian Ocean earthquake, both with near strike-parallel incidence.

Surface waves with near strike-parallel incidence ($\gamma = 0$) on vertical strike-slip faults also commonly show tremor modulation in phase with the Rayleigh wave following initial, energetic tremor bursts triggered by the Love wave. In this case as well, Rayleigh wave potential for triggering strike-parallel slip is minimal for strike-parallel incidence on vertical faults (see **Figure 7(a)**). Both observations suggest a possible role for the Rayleigh wave dilatational stress, which, although small, is the largest Rayleigh wave stress component for strike-parallel incidence on vertical faults.

Sufficiently energetic body waves are also capable of triggering tremor. Ghosh et al. (2009) found evidence for tremor on the Parkfield section of the San Andreas Fault triggered by the body-wave phases PKP, PP, SKP, and SKS from the $M_w = 9.1$ Sumatra earthquake. Hill et al. (2013) showed that the 100 s period S- and SS-waves from the $M_w = 9.0$ Tohoku-Oki earthquake also triggered tremor on the Parkfield

Table 3 Reported instances of dynamic triggering of tectonic tremor

Site	*Number of responses*	*Distance (km)*	*Triggering phase*	*Back azimuth (degrees)*	*Triggering stress (kPa)*	*References*
Subduction zones						
Nankai	6	1033–5100	L, R	Various	~10–20	1–5, 9
Cascadia	5	2309–12 400	L, R	134–311	~10	6–9, 10
S. Alaska, Aleutian	1	4232–5407	L, R	267–276	~10	9
Mexico	1	6751	S, L, R, O	–	5.7	11
Costa Rica	1	12 736	L	320.5	6	9
Southern Chile	1	15 067	R		1–10	17
New Zealand	2	9100–9259	R	S-N; 334	9–10	12, 9
Transform faults						
San Andreas	17		S, L, R, O	Various	0.7–10	9, 13, 14, 18, 19, 20
S. California	1	8606	L, R	307	9	9
Queen Charlotte	3–5		L, R	S-P	7	15
Oriente, Cuba	4	6203–12 479	L, R	80, 302, 328	5–10	10, 16
Garni, Armenia	1	7950	R	55	24	10
Other						
Central Taiwan	T	2652	S, L, R	46	~40	9, 10

Phases: L, Love; R, Rayleigh; S, S-wave; P, P-wave; O, others.
Back azimuth: S-P, strike parallel; S-N, strike normal.
References: 1. Obara (2003); 2. Miyazawa and Mori (2005); 3. Miyazawa and Mori (2006); 4. Miyazawa and Brodsky (2008); 5. Miyazawa et al. (2008); 6. Rubinstein et al. (2007); 7. Rubenstein et al. (2009); 8. Gomberg (2010); 9. Chao et al. (2013); 10. Gonzales-Huizar et al. (2012); 11. Zigone et al. (2012); 12. Fry et al. (2011); 13. Hill et al. (2013); 14. Peng et al. (2010a); 15. Aiken et al. (2013a); 16. Peng et al. (2013); 17. Peng et al. (2012); 18. Guilhem et al. (2010); 19. Shelly et al. (2011); 20. Guilhem and Nadeau (2012).

section of the San Andreas Fault. Zigone et al. (2012) documented triggering by S, ScS, and SS and both Love and Rayleigh waves on the Guerrero segment of the subducting plate beneath the southwest coast of Mexico by the 2010 $M_w = 8.8$ Maule earthquake. Miyazawa (2012) described evidence that the P-wave from the Tohoku-Oki earthquake may have triggered tremor bursts and brittle-failure earthquakes in volcanic and geothermal areas as it propagated to the southwest across Shikoku and the Nankai subduction zone.

Episodes of tectonic tremor are commonly triggered by teleseismic seismic waves with periods from ~20 to ~200 s and peak dynamic stresses on the order of ~10 kPa with individual tremor bursts modulated by stresses as small as ~1 kPa. Tidal stresses in the solid Earth with periods of 12 and 24 h also modulate tremor activity on both megathrust and transform plate-boundary faults (Hawthorne and Rubin, 2010; Lambert et al., 2009; Nakata et al., 2008; Shelly et al., 2007; Thomas et al., 2009, 2012). The tidal stresses are comparable with the minimum surface-wave stresses that trigger tremor, consistent with tremor-generating sections of the faults being extremely weak with near-lithostatic pore pressures. In the case of tidal-modulated tremor along the Parkfield sections of the San Andreas Fault, for example, the strike-parallel shear stress is ~0.1 kPa consistent with an effective fault-normal stress of between 1 and 10 kPa and a near-lithostatic pore pressure (Thomas et al., 2009, 2012).

4.11.4 Triggered Response Characteristics

Many of the response characteristics of tectonic tremor are covered in the previous section. Here, we focus primarily on the varied response characteristics of triggered earthquakes in the brittle domain.

4.11.4.1 The Global Search for a Triggering Threshold

Since the early evidence of dynamic triggering emerged, the seismological community has examined waveform characteristics from large remote events responsible for dynamic earthquake triggering for evidence of a triggering threshold. Threshold parameters explored include dependence on peak dynamic strain or stress, frequency, duration, and particle motion of the responsible seismic waves. Evidence for a clear threshold has yet to emerge; indeed, it seems that thresholds are likely site- and time-dependent.

Seismic-wave amplitudes responsible for dynamic triggering are variously reported as peak ground velocity; peak dynamic stress, T_p; or peak dynamic strain, ε_p (e.g., see note 20 in Hill et al., 1993). Although peak dynamic stresses at depth will in general differ from those based on seismograms recorded on the Earth's surface, they can be estimated given a reasonable model for the physical properties of the underlying crust (Gomberg, 1996). In computations of stress amplitudes at depth, for example, the tendency for surface-wave displacement amplitudes and strains to decrease with depth will be offset by one degree or another by the tendency of elastic moduli to increase with depth. Furthermore, surface-wave stresses show considerable variation over depths less than a wavelength (e.g., ~70 km for 20 s waves) depending on fault orientation, incidence angle, and velocity structure (see Figure 7).

As more studies of dynamic triggering have been conducted over the past decade, estimates of the minimum peak dynamic

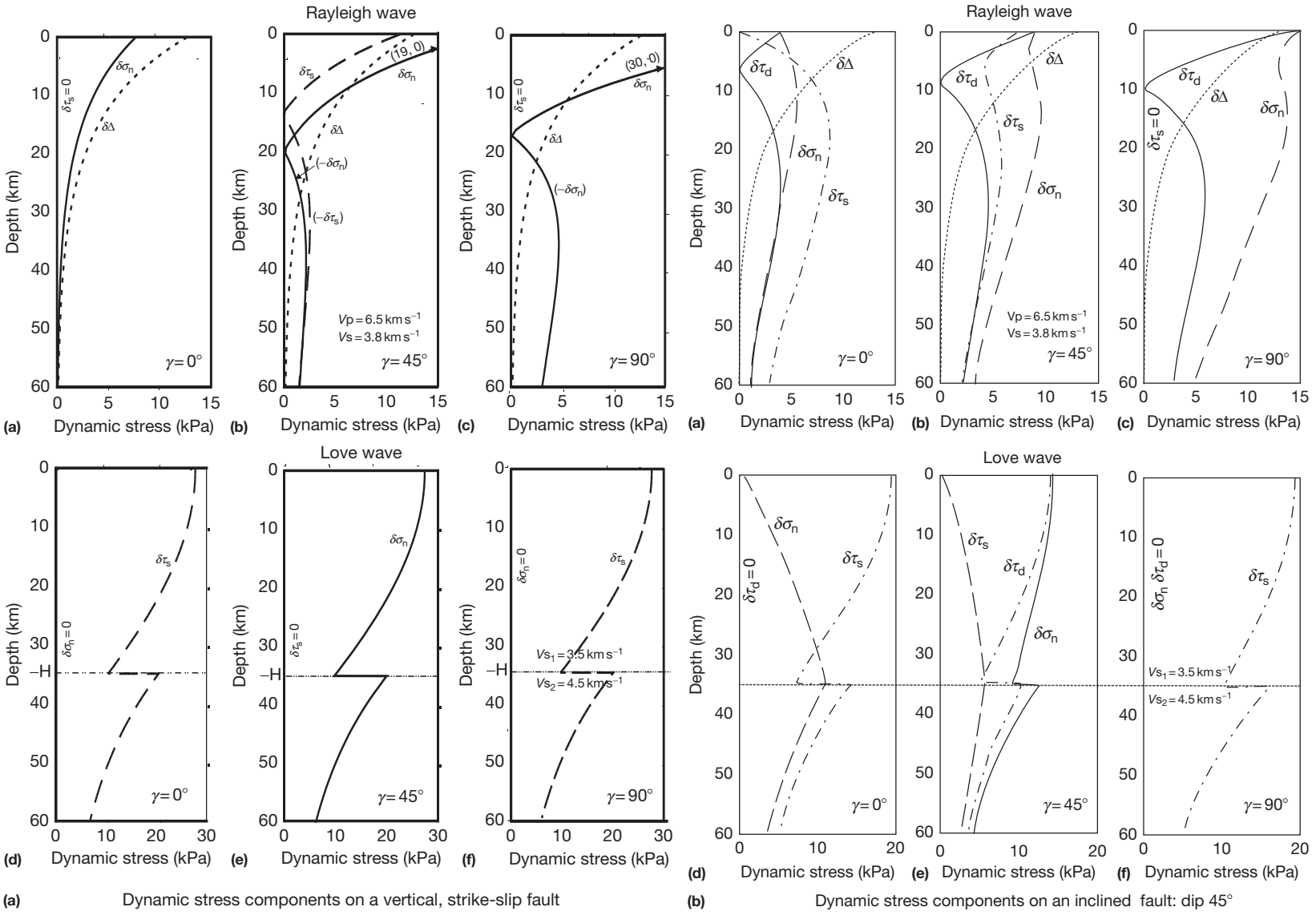


Figure 7 Peak values of surface-wave stress components as a function of depth for surface waves with period 20 s and displacement amplitudes of 1 cm incident on (a) a vertical fault and (b) inclined fault dipping 45°. Incidence angles are $\gamma = 0°$ (strike parallel), 45° (oblique), and 90° (strike normal). Stress components resolved on the fault planes are $\delta\sigma_n$, fault-normal stress; $\delta\tau_d$ and $\delta\tau_s$, dip-parallel and strike-parallel stresses, respectively; and $\delta\Delta$, the Rayleigh wave dilatational stress. *V*p and *V*s are the P- and S-wave velocities for the Rayleigh wave half-space model. Vs_1 and Vs_2 are S-wave velocities for the top layer and underlying half-space model for Love waves.

strain necessary to trigger earthquakes have been decreasing. Peak dynamic strain associated with specific remotely triggered earthquakes has been reported to be as small as 3×10^{-7} (~10 kPa) at the Coso volcanic field 3660 km from the 2002 $M_w = 7.9$ Denali Fault earthquake (Prejean et al., 2004). A recent study based on statistical analyses of earthquake catalogs suggests that dynamic earthquake triggering may result from peak dynamic strains as small as 3×10^{-9} or peak stresses of ~0.1 kPa (van der Elst and Brodsky, 2010). At the upper end, peak dynamic stresses can easily exceed 4 MPa within the transitional region to the aftershock zone and near field of a large earthquake (Kilb et al., 2002). Van der Elst and Brodsky (2010) suggested that a simple amplitude threshold may exist for specific areas in California and Japan based on their finding that the number of triggered earthquakes is linearly proportional to peak dynamic strain from large remote earthquakes.

The large range in peak dynamic stresses (or strains) that have resulted in remote triggering together with variations in intrinsic site characteristics indicate that the triggering process does not depend on a simple minimum amplitude threshold for dynamic stresses to be effective in all locales. Laboratory work by Savage and Morone (2008) suggests that variations in fault properties affect the potential for earthquake triggering. Thus, although most instances of dynamic triggering involve dynamic strains, $\varepsilon_p \sim 10^{-6}$, or dynamic stresses, $T_p \sim 30$ kPa (Gomberg and Johnson, 2005), this is neither a necessary nor a sufficient threshold for dynamic triggering (most areas with $T_p \sim 30$ kPa are not triggered and some areas with $T_p < 1$ kPa are triggered). In the case of the propagating front of triggered crustal earthquakes that moved across Japan with the P- and surface-wave velocities, Miyazawa (2011) concluded that the triggering P-wave strains of $\sim 10^{-7}$–10^{-8} (stresses ~0.3 to ~3 kPa) were an order of magnitude smaller than the surface-wave strains of $\sim 10^{-6}$ (stress ~30 kPa) from the 2011 Tohoku-Oki earthquake required for dynamic triggering.

The frequency of seismic waves may also be a factor in dynamic triggering thresholds in some environments. Specifically, dynamic stresses with periods of 20–30 s or longer commonly appear to be more effective at inducing a triggered response than those at higher frequencies. Brodsky and Prejean (2005) found that in a volcanic and geothermal environment, the Long Valley Caldera area lower-frequency seismic waves are more likely to trigger seismicity than higher-frequency seismic waves, as lower frequencies are more effective in stimulating fluid flow and fracture unclogging processes in hydrothermal systems (Brodsky et al., 2003). This is consistent with Parsons and Velasco (2009), who concluded that the lack of dynamic triggering at near-to-intermediate distances from underground nuclear explosions reflects minimal surface-wave amplitudes at periods ≥20 s generated by explosions. In contrast, Wu et al. (2012) showed that in a stable continental interior environment near Beijing, China, a triggering threshold for the area does not depend on frequency of shaking. Gomberg and Johnson (2005) also cited evidence for a frequency-independent threshold. Thus, it seems that a frequency threshold may only exist for specific sites, fault zone properties (Savage and Morone, 2008), or physical triggering mechanisms.

4.11.4.2 Aftershocks and Stress Triggering at Near-to-Intermediate Distance

Stress triggering at near and intermediate distances from a large earthquake is tied to the long-standing question of aftershock generation (Gomberg et al., 2003; Kilb et al., 2002; Nur and Booker, 1972; Marsan, 2003). The most intense aftershock activity generally occurs within a source dimension of the mainshock rupture (the near field), where static and dynamic stresses have comparable amplitudes and minimal temporal separation (**Figure 1**). Distal portions of an aftershock zone, however, commonly extend one to two source dimensions beyond the rupture zone in the range where static and dynamic stresses begin to assume separate identities. Efforts to distinguish between static stress triggering and dynamic stress triggering in aftershock zones have sought evidence in spatial distribution of aftershocks with respect to stress shadows predicted by static stress transfer models (e.g., Felzer and Brodsky, 2005; Kilb et al., 2002; **Chapter 4.10**, this volume; Richards-Dinger et al., 2010; Toda et al., 2012; Voisin et al., 2004) as well as for the influence of rupture directivity (Gomberg et al., 2003) and the variation in peak dynamic stress amplitudes with distance (Felzer and Brodsky, 2006). For example, Pollitz and Johnston (2006) pointed to the relative absence of aftershocks following slow earthquakes along the central section of the San Andreas Fault compared with typical earthquakes of similar location and seismic moment as evidence for the importance of dynamic stresses in aftershock production. Di Carli et al. (2008) found that the largest aftershock ($M_s = 5.5$) at a distance of 20 km from the 2000 $M = 6.7$ western Tottori (Japan) mainshock epicenter was located in the stress shadow of the mainshock and thus that it was dynamically triggered. Similarly, Meng et al. (2013b) found clear evidence that seismic waves from the 2010 $M_w = 7.2$ El Mayor–Cucapah earthquake produced dynamic triggering in the Salton Sea area in southern California, which lies within the static stress shadow of the El Mayor–Cucapah mainshock. Based on such studies, it seems that dynamic stresses are important in aftershock generation at all distances, particularly for short times (hours) following the mainshock.

The 2011 $M_w = 9.0$ Tohoku-Oki earthquake triggered seismicity throughout much of Japan, most of which was concentrated in the aftershock zone. Toda et al. (2011) cited evidence that static triggering of aftershocks in central Japan extended to distances of ~425 km from the Tohoku-Oki epicenter, and as Miyazawa (2011, 2012) documented, the propagation of triggered seismicity across southwestern Japan coincident with both the P-wave and surface waves. The surface wave-triggered activity extended ~1400 km southwest of the Tohoku-Oki epicenter with little evidence of propagation to the northeast suggesting the influence of rupture directivity and, therefore, dynamic stress.

This is currently a rapidly evolving research topic with the relative merits of static versus dynamic aftershock triggering the subject of energetic exchanges between advocates of one or the other viewpoint (Harris and Day, 1993; Harris et al., 2002; Stein, 1999; Voisin et al., 2004). Accumulating evidence suggests that dynamic triggering does indeed have a role in aftershock generation and in complex rupture processes in

which major earthquakes consist of multiple, large subevents (e.g., Aagaard et al., 2004; Antonioli et al., 2002; Rybicki et al., 1985; Voisin et al., 2000). Accumulating results also provide clear evidence for reduced seismicity in the static stress shadow quadrants of mainshocks (Meng et al., 2013a; Toda et al., 2012). Together, these results indicate that both static triggering and dynamic triggering contribute near-field aftershock generation.

4.11.4.3 Tectonic Setting and Source Radiation Patterns

The tectonic setting and stress regime have a first-order effect on the style of earthquake activity in any given area. Accumulating evidence indicates that these environmental factors also have an influence on patterns of remotely triggered seismicity. Although remote dynamic triggering occurs in all tectonic environments (Gonzalez-Huizar et al., 2012; Velasco et al., 2008), extensional or transtensional tectonic regimes, geothermal areas and areas of Quaternary to Recent volcanism, and the tremor-producing transition zones on major plate-boundary faults appear to be particularly susceptible. Dynamic triggering commonly coincides with areas of elevated background seismicity consistent with these areas being in a near-critical stress state.

Seismic-wave radiation patterns from a triggering mainshock can influence the spatial pattern of sites likely to respond to dynamic triggering. Love waves generally have a higher triggering potential than Rayleigh waves (Hill, 2012a,b). For a given moment and source depth, the four-lobe Love wave radiation pattern for strike-slip faults is more energetic than the corresponding pattern for Rayleigh waves. Peaks in the Love wave radiation pattern are in strike-parallel and strike-normal directions, while peaks in the Rayleigh wave pattern are rotated 45° with respect to the Love wave pattern. As a result, some of the most common and most compelling displays of remote dynamic triggering are associated with $M > 6.5$ earthquakes on strike-slip faults with unilateral rupture and pronounced strike-parallel directivity as exemplified by the $M_w = 7.4$ Landers, $M_w = 7.9$ Denali Fault, and the $M_w = 8.9$ Indian Ocean earthquakes.

In contrast, the two-lobe Rayleigh wave radiation pattern for dip-slip faults is more energetic than the corresponding four-lobe Love wave pattern for a given moment and source depth. Peaks in the two-lobe Rayleigh wave radiation pattern are in the slip direction and normal to fault strike (see Figure 11.19 in Dahlen and Tromp, 1998). Moment matters, of course, and the energetic Love and Rayleigh waves produced by megathrust earthquakes on subduction zones both contribute to widespread dynamic triggering as underscored by the 2004 $M_w = 9.1$ Sumatra–Andaman earthquake (Lay et al., 2005), the 2010 $M_w = 8.8$ Maule earthquake, and the 2011 $M_w = 9.0$ Tohoku-Oki earthquake. The lack of published descriptions of remote dynamic triggering attributable to large normal fault earthquakes may be largely a matter of sampling. Few $M > 6.5$ normal earthquakes have occurred in well-instrumented continental environments in the last 25 years. In any case, and in spite of triggering-potential expatiations based on theoretical seismic wave radiation patterns, there is no substitute for using observed seismic waveforms recorded on broadband seismic stations (if available) to evaluate the triggering potential of seismic waves incident on a nearby fault of known orientation.

The predominate remote response triggered by the Landers, Hector Mine, and Denali Fault strike-slip earthquakes was in the direction of maximum rupture directivity and in areas of transtensional tectonics characterized by elevated background seismicity (Hill and Prejean, 2007). This was underscored, for example, by the absence of a triggered response along the creeping section of the San Andreas Fault through central California and Parkfield that, although a persistent source of small earthquakes and in the peak directivity direction for the Landers rupture, is a transpressional stress regime (Spudich et al., 1995). In all three cases, many of the triggered sites were areas of geothermal activity and/or Quaternary volcanism.

In contrast, Jiang et al. (2010) found that the 2008 $M_w = 7.9$ Wenchuan earthquake in southern China triggered seismicity increases in the vicinity of widely scattered, historically active faults north of the epicenter (in the general direction of rupture propagation). The triggered seismicity occurred in both transtensional and transpressional tectonic regimes and mostly in areas not associated with geothermal or volcanic activity.

Extensional and transtensional stress regimes have three closely related properties that might enhance their susceptibility to dynamic triggering: (1) The least principal stress is subhorizontal. Extension cracks, which form in planes normal to the least principal stress, that are vertical. This facilitates the upward migration of crustal fluids from warmer, higher-pressure conditions at depth toward the surface through vertically oriented cracks. Volcanic and geothermal areas tend to be concentrated in extensional regimes. (2) As illustrated in **Figure 8(a)**, faults in an extensional regime are intrinsically weaker than those in compressional regimes given a uniform coefficient of friction and Andersonian faulting (Sibson, 1982). Hough and Kanamori (2002), for example, suggested that elevated pore pressures and temperatures associated with geothermal areas should enhance this difference (illustrated by the dashed Mohr's circle in **Figure 8(a)**). (3) Normal faults are more susceptible to triggering by surface waves than reverse faults in the upper crust (e.g., $z < 10$ km for 20 s waves), although this difference diminishes with increasing depth (Hill, 2012a).

Subduction-related volcanic centers are located on convergent plate boundaries. Locally, however, they commonly coincide with a zone of crustal extension and transtensional deformation associated with upward flexure of the overriding plate. This is the case for the Katmai volcanic field in Alaska (Moran, 2003), for example, and most of the Cascade volcanoes in the Pacific Northwest states. Mt. Rainier in Washington is exceptional with its setting in a transpressional stress regime. Even here, however, focal mechanisms for earthquakes occurring within the edifice of the volcano or shallow crust immediately below the volcano show a combination of strike-slip and normal faulting (Moran et al., 2000; Shelly et al., 2013).

Examples of evidence for remote triggering in compressional/transpressional regimes include the Bowen Basin in the stable Australian craton (Gonzalez-Huizar and Velasco, 2011), the Ohio River Valley area, and Atlantic coast states following the 1811–12 New Madrid, Missouri, earthquakes and the Charleston, South Carolina, earthquake in 1887 (Hough, 2001; Hough et al., 2003). Another involves evidence for apparent seismicity increases at distances of 70–200 km

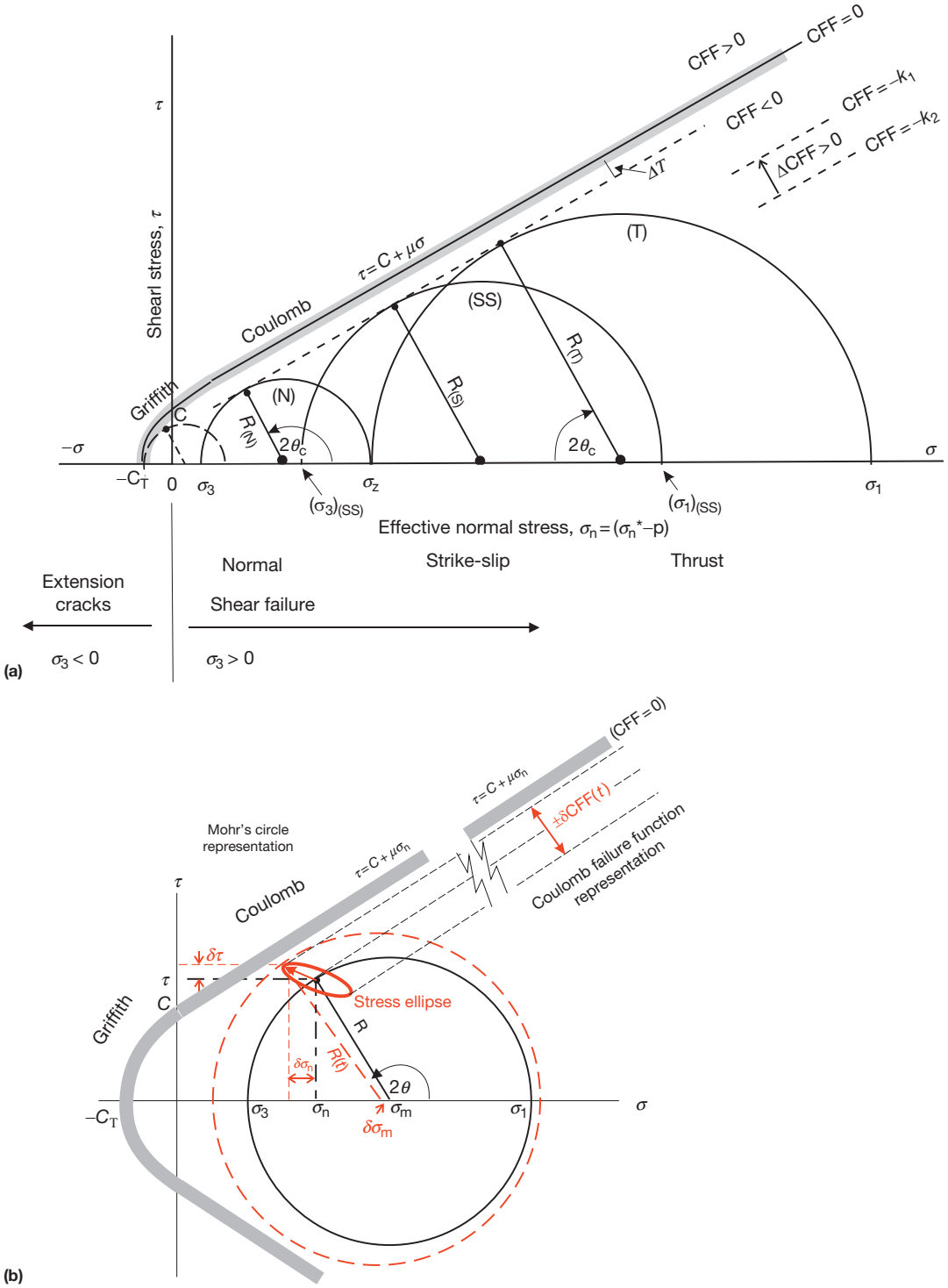


Figure 8 (a) Mohr's diagram comparing maximum stress levels for extensional (normal faulting, *N*) and compressional (thrust faulting, *T*) regimes in a pervasively faulted crust with a common (e.g., hydrostatic) pore pressure and with frictional strength limited by the Coulomb failure criteria. τ and σ are the shear stress and effective normal stress components, respectively, acting on a fault plane at an angle, θ, with respect to the least principal stress, σ_3. The coefficient of static friction is $\mu \sim 0.6$ with the gray band representing scatter in data. C and C_t are the cohesive and tensile strengths, respectively. The effective normal stress $\sigma = (\sigma_n^* - P)$, or the rock matrix normal stress σ_n^* reduced by the pore pressure, P. The vertical (lithostatic) stress, σ_z, is the greatest principal stress, σ_1, and least principal stress, σ_3, for the extensional (normal faulting) and compressional (thrust faulting) regimes, respectively. The radius, R, is half the stress difference, or $R = \frac{1}{2}(\sigma_1 - \sigma_3)$. Failure occurs when the Mohr's circles for optimally oriented faults ($\theta = \theta_c$) first touch the Coulomb failure envelope (CFF = 0). The strike-slip stress (SS) regime is intermediate between the thrust (T) and normal (N) regimes. The small dashed circle illustrates the case for superhydrostatic pore pressure in an extensional regime. Hydrofracturing results along vertically oriented cracks in the extensional regime if the pore pressure exceeds the least principal stress such that $\sigma_3 \leq 0$. (b) A dynamic stress orbit (red ellipse) in the $\sigma - \tau$ coordinate space of a two-dimensional Mohr's circle and the Coulomb–Griffith failure criteria. Black lines define the background stress state, $R(t < t_0)$, for a fault at an angle, $\theta(t < t_0)$, with respect to the least principal stress, σ_3. $\sigma_m = \frac{1}{2}(\sigma_1 - \sigma_3)$ is the two-dimensional mean stress. The red lines represent a snapshot in time of the stress state perturbed by dynamic stresses. The dynamic stress orbit is traced by the point R on the circumference of the circle through one stressing cycle. The dynamic Coulomb failure function δCFF(t) corresponds to the projection of the stress orbit onto a normal to the Coulomb failure envelope, $\tau = C + \mu\sigma_n$.

triggered by dynamic stresses associated with supercritical SmS waves from 15 moderate earthquakes ($M \sim 5$) in coastal California (Hough, 2005). In this case, several areas of slightly elevated β-statistic values were located within the transpressional Coast Ranges. Hough et al. (2005) also suggested that SmS waves may have triggered a $M > 7$ earthquake a few minutes after the $M_w = 7.8$ earthquake in 1905 in the compressional Kangra region of India.

This raises a question regarding the meaning of the crust being in a state of incipient failure nearly everywhere (Zoback and Zoback, 2002). Within the context of **Figure 8(a)**, a weak, extensional crust offers no particular advantage over a strong compressional crust if incipient failure means that, on average, the stress level everywhere hovers below the failure strength by some common increment, ΔT (this corresponds to a stress state defined by a line parallel with but incrementally below the Coulomb failure curve, CFF = 0 in **Figure 8(a)**). A weak crust would be more susceptible to dynamic triggering, however, if average stress levels were some large fraction of the failure stress (represented by a line with a smaller slope than that for CFF = 0 in **Figure 8(a)**). Currently available data seem ambiguous on this point. Any diminished propensity for triggering in stable intraplate regions due to a stronger crust may be at least partially offset, however, by the fact that seismic wave amplitudes generally decay more rapidly with distance in tectonically active regimes than in stable intraplate regions because of the higher intrinsic attenuation and the presence of velocity inversions in the lower crust and upper mantle that partially channel both body-wave energy and normal-mode (surface-wave) energy within depth intervals below the brittle crust (Bakun and McGarr, 2002).

4.11.4.4 Onsets, Delay Times, Durations, and Repeat Intervals

Rapid-onset triggering commonly coincides with the arrival of the large-amplitude Love or Rayleigh waves. Less commonly, it coincides with the S-wave arrivals (Hill et al., 2013; Prejean et al., 2004; Shelly et al., 2011; Zigone et al., 2012) and occasionally with P-wave arrivals (Ghosh et al., 2009; Miyazawa, 2012).

Reported delay times, Δt, between the arrival of the dynamic waves and the apparent onset of locally triggered seismicity vary from minutes to weeks or more. Delay times for the well-recorded instances of dynamic triggering from the Landers, Hector Mine, and Denali Fault earthquakes ranged from tens of seconds to between 24 and 33 h. The triggered response of the Western Ghats geothermal area in India to the 2001 $M_w = 7.6$ Bhuj earthquake was delayed by 2.5 h (Surve and Mohan, 2012). Delay times reported for suspected dynamically triggered earthquakes in the Taiwan region approach 15 days (Beresnev and Wen, 1995; Wen et al., 1996) and, in the Valley of Mexico, ~30 days (Singh et al., 1998). The longer the delay time, of course, the more tenuous the case for a causal link between the dynamic stresses and a local seismicity rate increase. Establishing a causal link between passing dynamic stresses and an isolated event or the onset of sequence with longer delay times requires independent, multiparameter data indicating the 'triggered' seismicity is a secondary response to some aseismic process (e.g., aseismic slip or pressure change in the case of volcanic systems) triggered by passing dynamic waves and/or a well-tested physical model that accounts for the delay between the arrival of dynamic stresses and the local onset of brittle failure.

Durand et al. (2010, 2013) described a pattern of rapid-onset and delayed triggering in the near-to-intermediate field following the 1999 $M_w = 7.6$ Izmit earthquake that they attribute to dynamic triggering and static triggering, respectively. The rapid-onset triggering was limited to strike-slip faults subparallel with the westward extension of the North Anatolian Fault in the eastern Marmara Sea west of the Izmit rupture and within a static stress shadow of the Izmit rupture. Here, the seismicity increased abruptly with passing of the seismic waves from the Izmit earthquake and abruptly died away. In contrast, seismicity in extensional areas to the north and south began more slowly and continued at elevated rates for months to years, which they attribute to static pressure changes and fluid flow.

Reported delay times based on recordings from standard, short-period seismic networks may overestimate actual delay times because local earthquakes are typically masked by the large-amplitude surface waves from large earthquakes. The development of broadband, high-dynamic-range instrumentation and digital processing has greatly enhanced our ability to recognize the early onset of local earthquake activity within coda waves from large, distant earthquakes through effective filtering. Prejean et al. (2004) found this to be the case in their search for dynamic triggering along the west coast of the United States following the $M_w = 7.9$ Denali Fault earthquake in 2002 (see **Figure 4**). Applying a β-statistic test to the Pacific Northwest and California earthquake catalogs produced no clear evidence for dynamic triggering. By applying a high-pass filter to seismograms from broadband and strong motion instruments, however, they discovered a plethora of small earthquakes embedded in the strong surface waves from the Denali earthquake at four sites in Washington and California from Mount Rainier ($\Delta \sim 3108$ km) south to the Coso volcanic field in southern California ($\Delta \sim 3660$ km). Small delay times were the rule for most instances of dynamic triggering from the Denali Fault earthquake.

This underscores two points: (1) Rapid-onset dynamic triggering with small delay times ($\Delta t \sim 0$) with respect to peak dynamic stresses is probably considerably more common than generally realized, and (2) our ability to detect dynamic triggering remains extremely uneven depending critically on the distribution of dense, high-quality seismic networks that include broadband, high-dynamic-range digital instruments.

Durations of dynamically triggered seismicity sequences range from a few minutes to several weeks. Short-lived episodes of triggered seismicity often appear as a rapid-fire sequence of overlapping, brittle-failure earthquakes (spasmodic bursts) within the surface-wave train. Longer-lived episodes more commonly evolve as earthquake swarms or foreshock–aftershock sequences with a temporal decay well described by a modified Omori's law of the form

$$n(t) = k/(t+c)^p \qquad [1]$$

where $n(t)$ is the number of earthquakes per unit time t and k, c, and p are parameters (Kisslinger and Jones, 1991). Pankow et al. (2004), for example, use this relation to estimate

durations of ~25 days for the seismicity triggered in central Utah by the 2002 Denali Fault earthquake. Hainzl and Ogata (2005) described a promising application of the epidemic-type aftershock sequence (ETAS) model that can be used to recognize the modulating role pore pressure diffusion in swarm-like triggered sequences. Generally speaking, longer durations tend to be associated with more energetic episodes of triggered seismicity. Indeed, Brodsky (2006) argued that the duration of most triggered sequences can in principle be predicted from the cumulative seismic moment of events triggered during the passage of seismic waves (dynamic stressing).

Many sites with documented remote triggering have responded more than once to the dynamic waves from distant earthquakes. Notable examples include the tremor-generating transition zones on the Nankai and Cascadia subduction megathrust faults, the Parkfield section of the San Andreas Fault, and the Central Range in Taiwan, which seem to respond to most $M > 7.5$ earthquakes. Other commonly repeating sites are in volcanic/geothermal areas including the Geysers geothermal field in northern California; the Coso volcanic field in southeastern California; the Katmai volcanic field in Alaska; Iwo Jima and Aso volcano in Japan; and the Fangshorn pluton in China. Peng et al. (2011a,b) described a special case of repeated triggering in which successive passes of surface waves circling the Earth (G_1, G_2, and G_3 and R_1 and R_2) from the 2010 $M_w = 8.8$ Chile earthquake triggered successive bursts of seismicity in the Coso geothermal field with delays of roughly 80, 160, and 240 min. Similar cases of tremor and microearthquakes triggered by multiple surface waves from the Chile mainshock were documented by Jay et al. (2012) and Zigone et al. (2012).

Repeating sites raise the question of recharge times. Each instance of remotely triggered seismicity releases locally stored energy, thus moving the responding site incrementally away from the near-critical state that existed just prior to the arrival of the dynamic stresses. The time required for the site to return to a near-critical state (the recharge time) will depend on a host of factors including (1) the energy released during the most recent episode of seismic activity whether remotely triggered or not and (2) the background rate of active processes feeding energy into the local crustal volume (Harrington and Brodsky, 2006; Miyazawa et al., 2005). As an example of (1), Rabaul caldera (Papua New Guinea) responded to a M7 earthquake at a distance of 180 km with a pronounced earthquake swarm but produced no detectable activity in response to a second M7 earthquake 2 months later at a distance of only 60 km (Mori et al., 1989). Wu et al. (2012) showed a similar need for recharge time for dynamic triggering in the Fangshorn pluton region near Beijing China. Recharge times for the Geysers geothermal area appear to be short – a matter of months or less (Gomberg and Davis, 1996). They are apparently much longer (but ill-constrained) for Long Valley Caldera where none of the triggered sites are strictly repetitive (Brodsky and Prejean, 2005). The short recharge times for triggered tectonic tremor on plate-boundary megathrust and transform faults are consistent with small stress drops, persistent background tremor, and exceptionally low fault strength ($\mu \leq 0.2$) such that the tremor-generating transition zone is in a perpetual state of incipient failure. Recharge times and their spatial–temporal fluctuations complicate efforts to establish triggering thresholds for dynamic stresses (e.g., Gomberg, 1996; Gomberg and Johnson, 2005). At the same time, they offer important clues to the processes behind dynamic triggering at any given site.

4.11.4.5 Triggering and Modulation by Solid Earth Tides

One of the long-standing puzzles in seismology has been why the periodic stressing of the Earth's crust by solid Earth tides appears to have little or no influence on temporal patterns of earthquake occurrence – particularly given evidence that broad sections of the crust are in a state of incipient failure. Exceptions appeared to be limited to (1) shallow earthquake swarm activity in volcanic and geothermal areas (Klein, 1976; Tolstoy et al., 2002; Wilcock, 2001) and (2) modulation of tectonic tremor in the transition zones along major plate-boundary faults (Ide, 2010; Rubinstein et al., 2008; Thomas et al., 2009, 2012). Tanaka et al. (2004) and Cochran et al. (2004), however, presented convincing evidence that the solid Earth tides in combination with ocean loading tides do indeed modulate the occurrence of crustal thrust earthquakes in convergent margins around the Pacific basin at the 10–20% level. The key to the latter apparently lies in correlating well-constrained focal mechanisms from a massive number of earthquakes (tens of thousands) with the phase of the tidal cycle that augment the local tectonic stress directions. These results are consistent with inferences based on laboratory measurements that cyclical stressing at the 1–4 kPa level should modulate ~1% of the background seismicity (Beeler and Lockner, 2003; Lockner and Beeler, 1999). While strictly speaking, solid Earth tides represent a quasistatic process (the inertial component in the equations of motion are negligible for solid Earth tides), they suggest a lower bound of 0.1–1 kPa as a low-frequency triggering threshold based on the strike-parallel, shear stress component of the tides capable of modulating tectonic tremor on the San Andreas Fault (Thomas et al., 2012).

4.11.4.6 Earthquake Hazard Implications of Dynamic Triggering

Most earthquakes triggered during passing seismic waves from large earthquakes are small to moderate ($\leq M_w = 5.0$–5.5, e.g., Anderson et al., 1994; Lei et al., 2011; Parsons and Velasco, 2011). Parsons et al. (2012) suggested that stress changes from surface waves of large, distant earthquakes are too short and spatially limited to trigger a damaging earthquake based on modeling stress changes from the seismic wave field on a fault. They based their modeling on 20–60 s surface waves recorded in Utah from the 2002 $M_w = 7.9$ Denali earthquake. Love waves from the $M_w = 9.1$ Tohoku earthquake recorded along the San Andreas Fault in California, however, had a period of 200 s (Hill et al., 2013). Longer periods imply coherent stresses over longer times and greater areas and thus the possibility of larger earthquakes triggered during dynamic stressing.

Accumulating evidence, however, indicates that larger, potentially damaging earthquakes are more likely to occur as a result of delayed triggering. This was underscored by the dramatic fivefold increase in the global rate of $M_w = 5.5$–7.0 earthquakes in the 6 days following the 2012 $M_w = 8.6$ Indian Ocean earthquake (Pollitz et al., 2012). Also, Gonzalez-Huizar et al. (2012) found that delayed dynamically triggered

earthquakes are larger than events triggered during the surface waves, and Lin (2012) suggested that surface waves from the 2008 $M_w = 6.6$ earthquake in Indonesia may have triggered a M = 6.9 earthquake in Hokkaido, Japan. More recently, it has been seen that the rate of $M > 5.5$ crustal earthquakes is significantly increased in the days following all large subduction thrust earthquakes (>M8.6) since 1960 by examining historical earthquake catalogs and moment tensors.

Although many questions remain over the significance of delayed triggering and the processes linking the triggering stress to the nucleation of the delayed earthquake, there can be little doubt that delayed triggering is real. The challenge lies in obtaining high-resolution, broadband data capable of resolving aseismic processes linking an applied dynamic stress to the nucleation of a triggered event, thereby precluding random coincidence (e.g., Tape et al., 2013). In any case, the chance that a local, damaging earthquake will be triggered by a major earthquake elsewhere on the globe, while finite, remains quite small.

4.11.5 Proposed Models

The rapidly accumulating evidence for widespread dynamic triggering has spawned a host of physical models to explain how relatively low-amplitude oscillatory seismic waves can trigger earthquakes at distances of 100s to 1000s of km. Here, we describe the more widely cited of these models and the conditions and locations under which each is viable.

Although most instances of instantaneous remote dynamic triggering coincide with surface waves from large earthquakes, a minority coincide with energetic body-wave arrivals (Ghosh et al., 2009; Hill et al., 2013; Miyazawa, 2012; Shelly et al., 2011; Zigone et al., 2012). Because dynamic stresses associated with both surface- and body-wave stresses are calculable as a function of depth, they provide a means of quantitatively evaluating the stress components acting on the source structure (most commonly a fault) hosting triggered activity. The same is true for tidal stresses responsible for modulating levels of crustal seismicity or tremor activity in the transitions zone (Thomas et al., 2009, 2012).

4.11.5.1 Triggering by Frictional Failure

This class of models involves direct triggering with either static or dynamic stresses providing the stress increment necessary to exceed the frictional strength of faults, thus leading to unstable slip and local earthquakes. This concept is commonly expressed in terms of a change in the Coulomb failure function, or

$$\Delta\mathrm{CFF} = \Delta\tau - \mu_s \Delta\sigma_n \qquad [2]$$

where

$$\mathrm{CFF} = |\tau| - \mu_s\sigma_n - C \qquad [3]$$

is the Coulomb failure function, τ and σ_n are the shear and normal stress components acting on the fault, respectively, μ_s is the static coefficient of friction, and C is the cohesive strength (e.g., Cocco and Rice, 2002; Harris, 1998; Kilb et al., 2002; Oppenheimer et al., 1988). CFF = 0 corresponds to Byerlee's law for frictional failure (Byerlee, 1980), and a positive change in CFF, or $\Delta\mathrm{CFF} > 0$, indicates that the stress state has moved incrementally toward CFF = 0 and Coulomb failure.

Under a generalized Coulomb failure model, shear failure on a fault is imitated when the shear stress, τ, acting on a fault plane exceeds the combined cohesive strength, C, and frictional strength ($\mu_s\sigma_n$) of the fault, or

$$\tau = \pm[C + \mu_s\sigma_n] \qquad [4]$$

such that the Coulomb failure function (eqn [3]) $\mathrm{CFF} \geq 0$. In the $\sigma-\tau$ space of a Mohr's diagram (Figure 8), the constants $\mathrm{CFF} = K_i$, form a family of straight lines of with slope, μ_s with CFF = 0 representing Byerlee's law for frictional failure. Slip evolution following the onset of shear failure may be governed by one of the several nonlinear laws such as the Dieterich–Ruina rate-state friction law (Dieterich, 1979, Chapter 4.04, this volume).

Frictional models are commonly discussed within the context of a steady loading rate (far-field plate motion or the steady extension of a spring in the case of a slider block model) and a 'clock change.' A 'clock advance,' for example, results if a small dynamic perturbation in the applied stress triggers a slip event that would not have otherwise occurred until the failure threshold was reached under the steady, far-field loading rate (e.g., Gomberg et al., 1997, 2005; Perfettini et al., 2003). One consequence of a clock advance in this context is that the time to the next slip event under far-field loading is inversely proportional to the far-field loading rate. Frictional models involve fluids only indirectly through the effect of pore pressure on the effective normal stress (or equivalently, an effective coefficient of friction) on fault planes. For dynamic stresses with periods of minutes or less, the rock matrix will behave as an undrained poroelastic medium with fluid transport having a negligible influence on pore pressure fluctuations (Cocco and Rice, 2002). Frictional models for dynamic triggering will be limited to areas where the stress state on preexisting crustal faults hovers below the Coulomb failure stress by less than the peak dynamic stress amplitude.

4.11.5.1.1 Coulomb failure under dynamic stresses

By analogy with eqn [2], the dynamic Coulomb failure function is

$$\delta\mathrm{CFF}(t) = |\delta\tau(t)| - \mu_s\delta\sigma_n(t) \qquad [5]$$

(Kilb et al., 2000). In the stable domain ($\mathrm{CFF} < 0$), a positive $\delta\mathrm{CFF}(t)$ indicates that the stress state has moved incrementally toward failure. Under the Coulomb failure model for dynamic triggering (Figure 9), the traction $\mathbf{T}[R(t), \theta(t)]$ on a fault at angle θ with respect to the least principal stress, σ_3, of the initial stress state, $\mathbf{T}(t_0)$, begins an oscillatory trajectory as the seismic waves arrive at time t_0. A triggered earthquake results if the trajectory of the stress orbit enters the Coulomb failure zone $\mathrm{CFF} \sim 0$ (gray band in Figure 9).

One can estimate an in situ triggering potential, $0 \leq P(\gamma) \leq 1$, for seismic surface waves as a function of incidence angle, γ, on a fault of known orientation based on the shape and orientation of the stress orbit for given wave type with respect to the Coulomb failure envelope with a friction coefficient appropriate for the target fault (Hill, 2012a,b). $P(\gamma)$ represents the maximum positive value of $\delta\mathrm{CFF}$ normalized by the largest semimajor axis, L, for the stress orbits spanning the entire

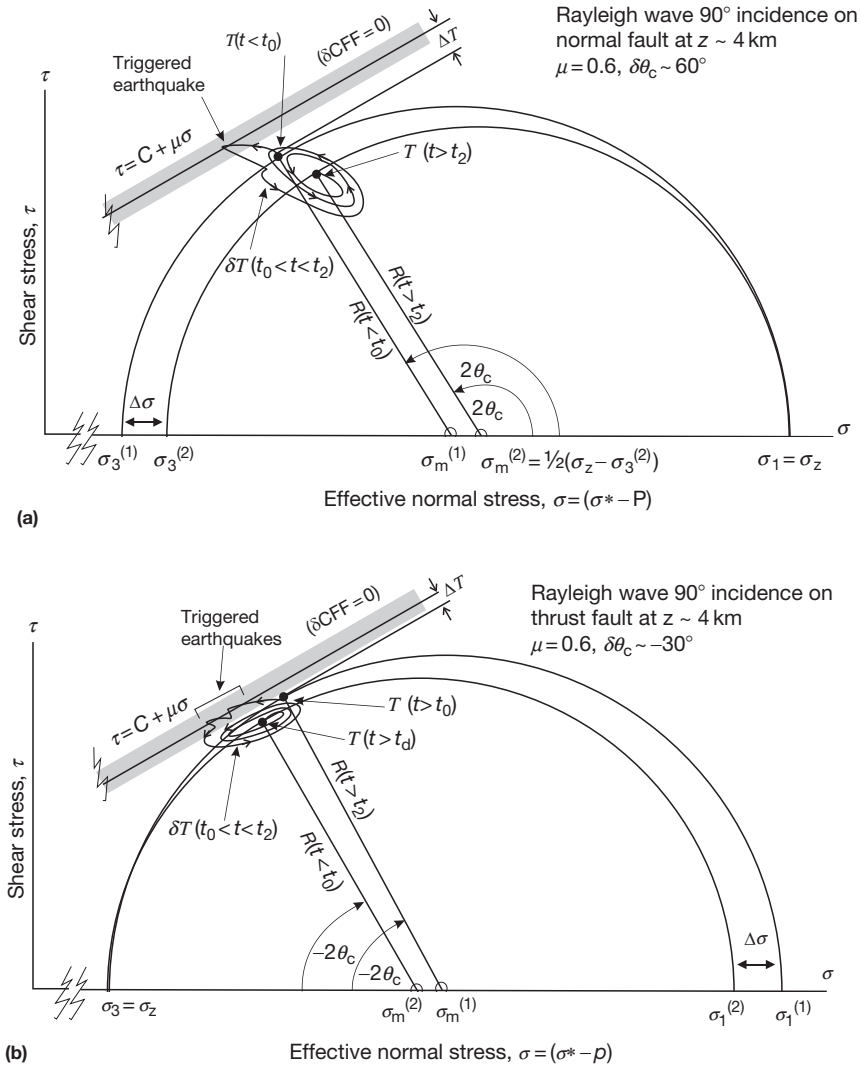


Figure 9 Schematic illustration comparing the dynamic triggering under the Coulomb failure model for (a) extensional and (b) compressional regimes. The point $\mathbf{T}(t<t_0)$ represents a stable stress state slightly below the failure threshold on a fault at an angle $\theta(t<t_0)$ with respect to the local least principal stress direction, σ_3, which is horizontal in extensional regimes (a) and vertical in compressional regimes (b). Note that the vertical stress (lithostatic load), σ_z, corresponds to σ_1 in (a) and σ_3 in (b). At $t=t_0$, dynamic stresses from a distant earthquake arrive producing an oscillatory perturbation $\delta\mathbf{T}(t_0<t<t_2)$ in both the orientation and the magnitude of the local stress field. Slip occurs whenever $\delta\mathbf{T}(t_0<t<t_2)$ intersects the Coulomb failure zone, CFF ~ 0 (gray band) confining $\mathbf{T}(R(t), \theta(t))$ excursions to within or below this zone (CFF $\geq$ 0). The dynamic stresses die away at time t_2, and the stress state on the fault settles to $\mathbf{T}(t>t_2)$ reflecting the stress drop associated with the triggered seismicity. The lithostatic load, σ_z, corresponds to the greatest and least principal stresses in extensional and compressional regimes, respectively.

range of incidence angles, $-90^\circ \leq \gamma \leq 90^\circ$, or $P(\gamma) = |\delta\text{CFF}|_m/L$ (see Figure 10 in Hill, 2012b). In the case of Love waves incident on a vertical fault, the stress orbits are rectilinear and of the same length for all incidence angles. They rotate in $\delta\sigma - \delta\tau$ space from 0 to 2π as the incidence angle varies from 0 to π. Thus, $P(\gamma)$ varies from 0 to 1 as the orbit rotates from a parallel to perpendicular orientation with respect to the slope of the Coulomb failure envelope (Figures 8 and 10 in Hill, 2012a). Note that (1) $P(\gamma)$ is a 'floating' potential in the sense that the absolute ambient stress acting on a given fault and thus its proximity to criticality are generally not known and (2) its value depends on the assumed friction coefficient, μ_s (the slope of the Coulomb failure envelope). Results for surface waves indicate that for most configurations, Love waves have a higher triggering potential than Rayleigh waves with comparable surface displacement amplitudes but that details depend critically on fault type (e.g., normal, reverse, and strike-slip), orientation, depth, friction coefficient, and the wave incidence angle, γ.

Figure 10 illustrates the sensitivity of triggering potentials to incidence angle, γ, and friction coefficient, μ, on vertical

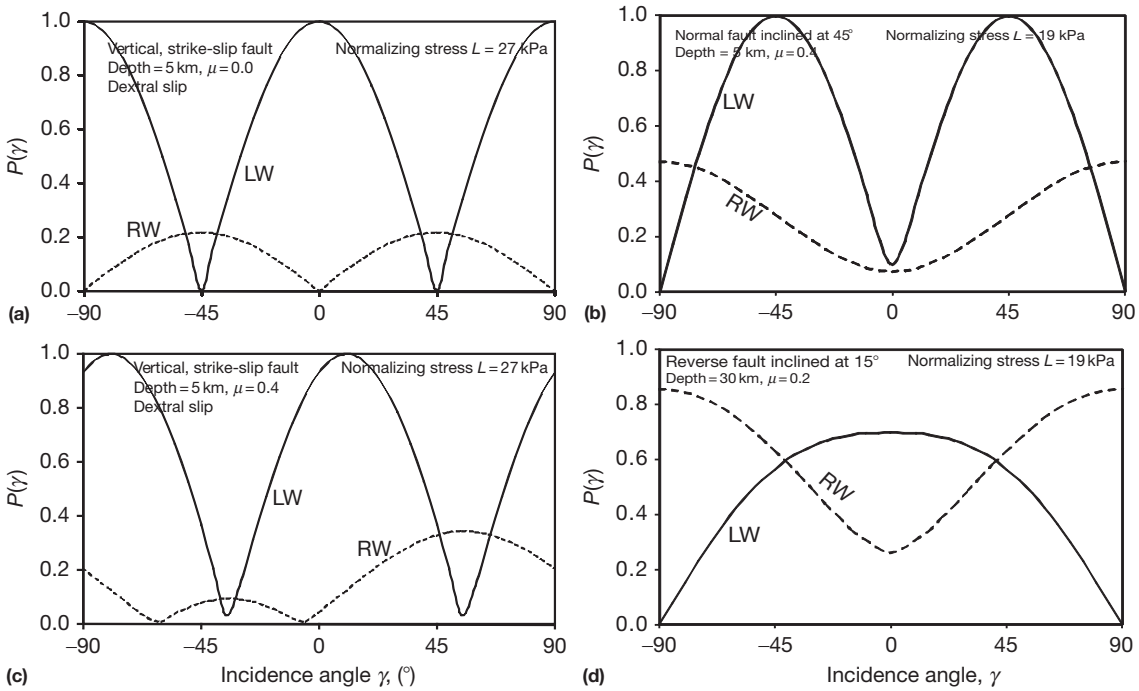


Figure 10 Variations in triggering potentials as a function of incidence angle, γ, for Love and Rayleigh waves incident on a vertical right-lateral, strike-slip fault at 5 km depth with friction coefficients (a) $\mu = 0$ and (b) $\mu = 0.4$ and inclined faults (c) with normal slip on a fault dipping 45° at a depth of 5 km and (d) reverse slip on a fault dipping 15° at 30 km with dip of 15°. The friction coefficients are $\mu = 0.4$ and 0.2 for (c) and (d), respectively. Multiplying the triggering potential $P(\gamma)$ by the normalizing stress, L, gives amplitudes of the strike-parallel (a) or dip-parallel (b) shear stress components acting on the respective fault planes for 20 s Love and Rayleigh waves with comparable displacement amplitudes, $A = 0.01$ m, at the Earth's surface.

strike-slip faults (Figure 10(a) and 10(b)) and dip-slip faults at two depths both inclined at 45° (Figure 10(c) and 10(d)). Note in the case of vertical strike-slip faults (left panels) the following:

- Love and Rayleigh wave potentials are anticorrelated with Love wave potentials dominating except near peaks in Rayleigh wave potentials.
- For a frictionless fault (Figure 10(a)), the potential curves are symmetric about $\gamma = 0$ (strike-parallel incidence) with Love wave potentials peaking at 0° and ±90° and Rayleigh wave potential peaking at ±45°.
- For a finite friction ($\mu = 0.4$ in this case), the peaks and troughs are shifted to the right for dextral slip and by an equal amount to the left for sinistral slip (not illustrated). The amount of shift varies with the friction coefficient.

Note in the case of the inclined, dip-slip faults the following:

- At 5 km depth and $\mu = 0.4$, the Love wave potential peaks at ±45°, while the Rayleigh wave potential for normal faults peaks at ±90°. The Love wave potential dominates except in the vicinity of ±90°. Both approach null values at 0°.
- At 30 km depth (tremor-generating depths) and $\mu = 0.2$, Love and Rayleigh wave potentials are anticorrelated with comparable peak amplitudes. The Love wave peaks at 0° and the Rayleigh wave potential for reverse faults peaks at ±90°.

The complex variations in triggering potentials with fault orientation, incidence angle, depth, and fault friction for the examples in Figures 7 and 10 emphasize the importance of having some constraints on these parameters when attempting to understand what role the various surface-wave stress components play in a given incidence of triggered seismicity.

As illustrated in Figure 9, the stress drop associated with the triggered earthquake alters the stress state (its orientation, θ, and the stress difference, $R(t_2) < R(t_0)$) in the vicinity of the fault such that the traction on the fault settles to a new value, $\mathbf{T}$ ($t > t_2$), after the seismic waves have passed. In the case of small stress drops and weak fault, bursts of triggered earthquakes may occur in phase with several cycles of the passing surface waves. In Figure 11(a), for example, small earthquakes triggered at Mt. Wrangell, Alaska, occur in phase with peaks in the horizontal extensional stresses during the Rayleigh wave train from the $M_w = 9.1$ Sumatra–Andaman Islands earthquake on 26 December 2004 (West et al., 2005). In Figure 11(b), tremor bursts on the Parkfield section of the San Andreas Fault occur in phase with the SH, SHSH, Love, and Rayleigh waves from the $M_w = 9.0$ Tohoku-Oki earthquake on 11 March 2011 (Hill et al., 2013). Multiple earthquakes may occur during a single dynamic stress cycle if (1) the initial triggered earthquake produces its own aftershock sequence, (2) the stress cycle triggers earthquakes on multiple nearby faults, or, (3) for the case of reverse slip on low-angle faults as illustrated in Figure 9(b), the semimajor axis of the stress orbit is subparallel with the

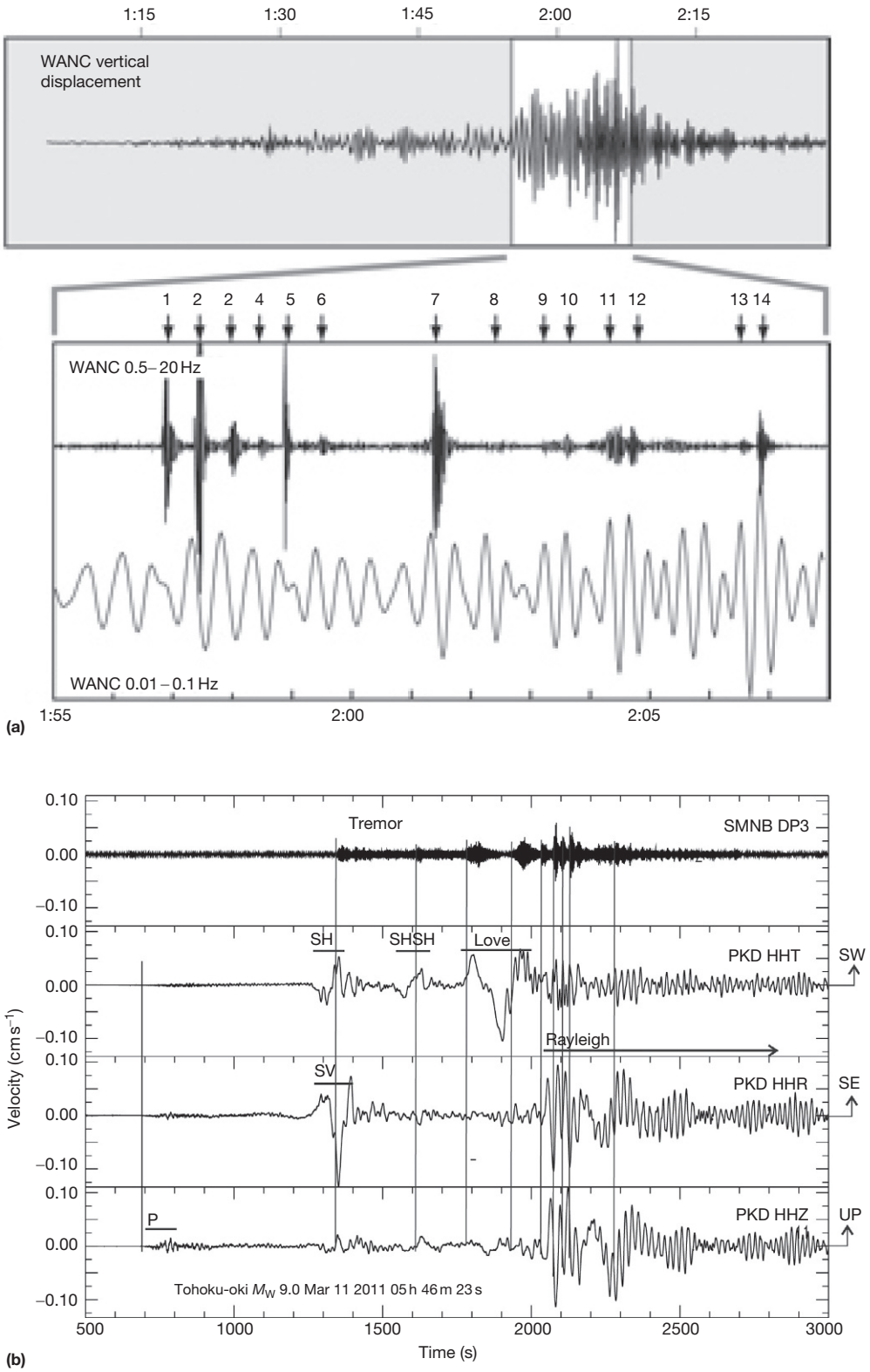


Figure 11 Examples illustrating the phase-lock between triggering teleseismic waves and triggered seismicity for (a) Mt. Wrangle, Alaska, seismicity triggered by the 2004 $M_w = 9.1$ Sumatra–Andaman earthquake. (Top) Short-period vertical displacement seismogram at the summit station WANC. Time in hours. (Bottom) Expanded view of highlighted section top panel showing triggered seismicity band-passed at 0.5–20 Hz and the Rayleigh wave coda band-passed at 0.01–0.1 Hz. Reproduced from West M, Sanchez JJ, and McNutt SR (2005) Periodically triggered seismicity at Mount Wrangell, Alaska, after the Sumatra earthquake. *Science* 308: 1144–1146, with permission. (b) Tectonic tremor on the Parkfield section of the San Andreas Fault triggered by the 2011 $M_w = 9.0$ Tohoku-Oki earthquake. (Top) Tremor recorded by short-period station SMNB. (Bottom) Three traces are teleseismic waveforms recorded on the broadband station PKD. Reproduced from Hill DP, Peng Z, Shelly DR, and Aiken C (2013) S-wave triggering of tremor beneath the Parkfield, CA, section of the San Andreas Fault by the 2011 Tohoku-Oki Japan earthquake: Observations and theory. *Bulletin of the Seismological Society of America* 103: 1541–1550. http://dx.doi.org/10.1785/0120120114, courtesy of the Bulletin of the Seismological Society of America.

Coulomb failure envelope such that a single cycle of the orbit intersects the failure envelope several times.

In principle, the Coulomb failure model should be viable in any region where faults are critically stressed and favorably oriented with respect to the polarity of the dynamic stresses from a distant earthquake. If most of the seismogenic crust is pervasively fractured and critically stressed as suggested by Zoback and Zoback (2002), earthquakes triggered by this process have an equal likelihood of occurring nearly everywhere. If this were the dominant triggering mechanism, one would expect that, given some minimum dynamic stress level, the distribution of observed triggered seismicity would correlate directly with the density of seismic networks and the sensitivity of earthquake detection techniques. As Rivera and Kanamori (2002) argued, however, the heterogeneity of both the stress field and the frictional strength seems to be an essential property of the crust. In a similar vein, Lapusta and Rice (2003) argued for heterogeneous strength along a given fault such that earthquake nucleation is confined to weak spots, while the average fault strength remains well above the Coulomb failure threshold.

4.11.5.1.2 Nonlinear friction

Extensive laboratory experiments on rock friction reveal that frictional failure involves a much richer range of (nonlinear) behaviors than predicted by the simple first-order model involving a constant static μ_s and dynamic μ_d coefficients of friction with $\mu_s > \mu_d$.

The Dieterich–Ruina rate-state friction laws (Dieterich, 1979, **Chapter 4.04**, this volume) have been widely used in studies of both static triggering and dynamic triggering. These rate-state laws, in which static friction retains a memory of sliding history along the fault surface and dynamic friction depends on the sliding velocity and an evolving state parameter, predict a range of behaviors depending on temporal variations in the applied stress and the parameter for a particular constitutive regime (Dieterich, 1979; Scholz, 1998). They include the possibility of delayed triggering (Parsons, 2005) and an Omori-like decay of triggered seismicity after shaking has stopped (Gomberg et al., 2005; Parsons, 2005). In the unstable, velocity-weakening regime, rate-state friction appears to be most effective under conditions of low effective normal stress (high pore pressure and weak faults) together with high-frequency vibrations (Gomberg et al., 2003; Perfettini et al., 2003; Scholz, 1998). Under rate-state friction in a conditionally stable (velocity-strengthening) regime, dynamic stresses can temporarily convert stable sliding (steady fault creep) to a stick–slip mode. The result is 'new seismicity' in the sense that the earthquakes generated during the stick–slip mode would not have occurred in the absence of dynamic triggering (Gomberg et al., 1997).

Subcritical crack growth or stress corrosion has long been known in material science to make cracks unstable (Anderson and Grew, 1977; Atkinson, 1984; Gomberg et al., 2001). When a crack experiences a change in its environment, such as a sudden increase in loading in the presence of particularly high temperatures and fluids, it may grow due to weakening at the crack tip by chemical corrosion. The crack may then rupture catastrophically. Thus, a sudden increase in loading or oscillatory loading (dynamic stressing) may shorten the time to earthquake failure. The equations governing subcritical crack growth are mathematically identical to those governing failure by rate-state friction (Kanamori and Brodsky, 2004), and thus, as Brodsky and Prejean (2005) pointed out, subcritical crack growth requires near-lithostatic pore pressures to be effective as a dynamic triggering process.

Daub et al. (2011) used both laboratory and field data to develop a brittle–ductile friction model that mimics tectonic tremor in the transition zones of major plate-boundary faults. They found that tremor-like signals result if 40–70% of the contact asperities on a fault interface are brittle (the remainder being ductile). They suggest that at tremor-generating depths of ~25 km, the brittle elements likely consist of feldspar, which remains brittle at temperatures of ~600 °C under low effective normal stresses consisted with a weak fault.

The response of granular media to imposed dynamic stresses offer insight on another possible mode of nonlinear frictional failure. Laboratory experiments imposing acoustic vibrations on granular media under low normal stress (~7 kPa) suggest that seismic waves incident on a weak fault with a core composed of fault gouge (the granular medium) with peak amplitudes comparable to the normal stress can result in an abrupt decrease in the modulus of the fault gouge and catastrophic slip in the form of an earthquake (Johnson and Jia, 2005; Johnson et al., 2008, 2012; van der Elst et al., 2012).

Because each of the nonlinear models (rate-state friction, granular media, and subcritical crack growth) mentioned earlier apparently requires near-lithostatic pore pressures to be effective in dynamic triggering (Brodsky and Prejean, 2005), it seems unlikely that they offer a universal explanation for dynamic triggering. Any of the three, however, may be important in volumes where pore pressures approach the lithostatic limit. All have pressure-sensitive failure thresholds such that the Coulomb failure model provides a useful criterion for the initiation of catastrophic slip. Nonlinear friction will take over once accelerated slip is under way governing subsequent behavior including sustained episodes of triggered earthquakes or tremor apparently modulated by the coda of teleseismic waves (see **Figure 11**).

4.11.5.2 Triggering Through Excitation of Crustal Fluids

Crustal fluids play a critical role in a wide range of tectonic and magmatic processes (see **Chapter 4.12**, this volume), including dynamic triggering. Large earthquakes have long been recognized as capable of disturbing hydrologic regimes at distances of thousands of kilometers beginning with the $M_w = 9.2$ 1964 Alaska earthquake (Roeloffs, 1996; Vohis, 1967), and fluids have long been known to play a role in earthquake rupture (e.g., Nur and Booker, 1972). These observations have inspired a class of physical models in which dynamic strains from the distant earthquake trigger local seismicity through fluid transport and pore pressure changes. This process may modify the Coulomb failure function (**Section 4.11.5.1.1**) such that the effective normal stress is decreased sufficiently to trigger failure (Beeler et al., 2000; Cocco and Rice, 2002) or that quasistatic (aseismic) strains associated with local, fluid-driven deformation are sufficient to trigger earthquakes. Fluids are active agents in geothermal

and volcanic areas, which appear to be particularly susceptible to dynamic triggering, and a number of explanations for triggered seismicity involving the movement of fluids have been proposed in the literature (e.g., Brodsky et al., 1998; Hill et al., 1995; Johnston et al., 1995; Linde et al., 1994; Manga and Brodsky, 2006; Moran et al., 2004). The hydrologic response to earthquakes includes changes in stream and spring discharge, groundwater properties (geochemistry, temperature, and turbidity), water-level fluctuations in wells, mud volcanoes, and soil liquefaction (**Chapter 4.12**, this volume; Wang and Manga, 2009). Wang and Manga (2009) showed that aside from liquefaction, which is limited to the near field, most hydrologic phenomena in the intermediate and far field are dynamically triggered. They suggest that changes in permeability resulting from oscillatory deformation and fluid flow associated with dynamic stresses play an important role in the triggering process.

In geothermal and volcanic areas, the simple linear version of 'clock advance' may be compromised because, in addition to steady loading from far-field tectonic stresses, the advection of magmatic fluids and heat from lower crustal or upper mantle source supplies local energy into a recently triggered volume of the brittle crust. As is evident from the behavior of the Long Valley Caldera volcanic–geothermal system over the past 25 years, these processes may operate episodically over timescales ranging from weeks to years (Hill, 2008). Long Valley Caldera responded with remote triggering multiple times during periods of active unrest and deformation in the 1990s. Since 2000, however, unrest has been minimal and the caldera has failed to respond with dynamic triggering by any of the four $M \sim 9$ earthquakes since 2000.

The susceptibility to dynamic triggering does not correlate with unrest in all volcanic systems, however. Japan with its many active volcanoes, for example, appears to be relatively insensitive to dynamic triggering (Harrington and Brodsky, 2006). With the exception of the Katmai volcanic cluster, Alaskan volcanoes show a similar insensitivity to dynamic triggering. The reason for this is not yet clear, although the compressional tectonic environment common to both areas may be a factor.

4.11.5.2.1 Hydrous fluid transport: changes in permeability and pore pressure

In this class of models for earthquake triggering, the dynamic strains from a distant earthquake modify crustal permeability by disrupting clogged fractures and hydraulic fracturing, thus leading to a redistribution of pore pressure (Elkhoury et al., 2011; Wang and Manga, 2010). These models should be viable in any area where isolated pockets of high-pressure fluids develop. Brodsky et al. (2003), for example, proposed that coseismic water-level changes observed in wells at Grants Pass in southern Oregon might result from dynamic stresses from the 1994 $M = 7.2$ Petrolia, CA, and the 1999 $M = 7.4$ Oaxaca, Mexico, earthquakes at ~300, ~3850 km, respectively, opening permeable fractures clogged by accumulating detritus over time.

The pore pressure redistribution mechanism may be particularly relevant in active geothermal areas, such as the Geysers and Coso geothermal fields, as fractures are sealed and high-pressure compartments form over relatively short timescales as minerals are precipitated from hot brines. The mechanism may also be important in the crust above a magma body in which hydrous magmatic fluids may reach near-lithostatic pressures in the low-permeability plastic zone between the roof of the magma body and the brittle crust, where hydrothermal fluids circulate under hydrostatic conditions. This is the hydraulic surge model proposed by Fournier (1999) in which dynamic stresses from a large distant earthquake may rupture this impermeable zone, permitting high-pressure fluids to surge into the overlying brittle crust. This mechanism may explain both earthquakes triggered during and after the wave train of the large distant earthquake has passed as pore fluids move by diffusion through the crust.

Hough and Kanamori (2002) looked for evidence of fluid movement in the failure process of earthquakes triggered in the Salton Sea by the 1999 Hector Mine earthquake by examining earthquake source parameters. They found that the triggered earthquakes had essentially normal to slightly low stress drops and spectral contents typical of brittle shear failure with no resemblance to the long-period (LP) earthquakes or tremor-like sequences often seen in volcanic areas (Chouet, 1992; Julian, 1994).

4.11.5.2.2 Magmatic fluids

Growing evidence indicates that dynamic stresses are capable of perturbing the state of crustal or upper mantle magma bodies triggering internal pressure changes and, in some cases, eruptions (Hill et al., 2002; Linde and Sacks, 1998; Manga and Brodsky, 2006; Marzocchi, 2002). Under these models, triggered seismicity is a secondary result of locally triggered deformation associated with the change in state of a nearby magma body or intrusion from a critically pressurized magma body. Because these models predict distinctive patterns of ground deformation associated with triggered seismicity, continuous, high-resolution deformation monitoring networks are key to testing their validity.

Bubble excitation by dynamic stressing may explain some earthquake–magmatic interactions. Bubbles play a central role in the source mechanisms for LP and very LP volcanic earthquakes (Chouet, 1992) and in eruption dynamics (Mangan and Sisson, 2000). Two highly idealized physical models widely mentioned as possible explanations for remotely triggered seismicity appeal to changes in bubble pressurization in a two-phase fluid. In the rectified diffusion model, volatiles in a saturated fluid are selectively pumped into bubbles during the dilatational phase of each strain cycle as the wave train of the distant earthquake passes (Brodsky et al., 1998; Sturtevant et al., 1996). In the idealized advective overpressure model, a bubble adhering to the walls of a magma body is shaken loose by the dynamic stresses from the remote earthquake. Under the assumptions of rigid walls and an incompressible magma, the bubble maintains the same size as it buoyantly rises a distance, h. By the gas law, the product of bubble volume and pressure is constant in an isothermal system. Thus, the pressure in the magma chamber increases by $\rho g h$, where ρ is magma density and g is the acceleration of gravity (Linde et al., 1994; Sahagian and Proussevitch, 1992). In both models, increasing pressure in the bubbles is transmitted to the surrounding interstitial fluid, thereby increasing pore pressure and triggering earthquakes. These models could apply to either hydrous fluids or

magma. Both, however, have been criticized on the basis that they depend on restrictive assumptions that are unrealistic under natural conditions in the Earth (Bagdassarov, 1994; Brodsky et al., 1998; Ichihara and Brodsky, 2006; Pyle and Pyle, 1995).

Bubble nucleation stimulated by dynamic stressing in a supersaturated magma is yet another fluid activation model (Manga and Brodsky, 2006). Under proper conditions, small pressure changes associated with dynamic stresses in a crystallizing magma that is close to critical supersaturation should be capable of triggering a significant excess in bubble nucleation leading to a marked pressure increase and deformation with local seismicity as a secondary response. Numerical modeling by Shimomura et al. (2006) and Chouet et al. (2006) indicates that when subject to a small externally imposed pressure drop, a densely packed matrix of tiny bubbles in a magma-filled crack can lead to rapid diffusion-driven bubble growth and volumetric expansion of the crack. These models have yet to be evaluated for their response to realistic dynamic stresses, but they keep open the possibility that bubbles may have an important role in the triggered response of magmatic systems to dynamic stresses.

Bubble excitation in a critically pressurized magma body may be sufficient to trigger an intrusion into the surround crust accompanied by seismicity in response to deformation associated with the intrusion. Dilatational strain meters at Mammoth Mountain recorded deformation transients coincident with the seismicity triggering in the Long Valley Caldera area following the Landers, Hector Mine, and Denali Fault earthquakes (**Figure 6**). Based on these data, Johnston et al. (2000) and Johnston et al. (2004a) proposed that, in the case of Mammoth Mountain and Long Valley, the dynamic waves from distant earthquakes trigger magmatic intrusions into the shallow crust. These intrusions, which may involve a massive hydrofracture by aqueous magmatic fluids or magma itself, then trigger earthquakes by changing the static stress field or locally increasing the pore pressure as fluids are exsolved. Thus, the local seismicity is a secondary response to the initiating ground motion.

Dynamic waves from a large distant earthquake may disrupt the stability of a partially crystallized magma body releasing accumulated deviatoric stress supported by the interconnected crystal structure (Hill et al., 1995). As in the magmatic intrusion model, the local seismicity triggered by a relaxing magma body is responding to a local, quasistatic strain source rather than the initiating dynamic stresses. The recharge time for this model is presumably rather long depending on both the crystallization rate in cooling magma body and the far-field strain rate.

Yet another possible response of a magma body to dynamic stresses suggests that passing seismic waves may dislodge dense crystals from the ceiling and walls of a magma chamber. As these crystals then sink due to gravity, they may stimulate convection in the magma chamber. Bubbles would presumably nucleate and grow in the ascending volatile-rich magma, thereby increasing pressure in the magma body and deforming the overlying crust (Hill et al., 2002). Manga and Brodsky (2006) showed that this process may be capable of creating sufficient overpressure to trigger a volcanic eruption on the timescale of days. It is likely the process could trigger earthquakes on a similar timescale.

4.11.6 Challenges for the Future

Because the in situ perturbing dynamic stresses are calculable, the study of remotely triggered seismicity gives us a unique window into the earthquake nucleation process. Although this rapidly advancing field has great potential for illuminating the earthquake initiation process, many fundamental issues remain regarding the detection, temporal and spatial distribution, and physical processes behind dynamically triggered seismicity. Some of these issues include the following:

- *Uniform detection of triggered seismicity*. Automated earthquake detection algorithms are ineffective at identifying small triggered events recorded by a single, low-dynamic-range station. Reliable identification and location of triggered seismicity buried in the surface-wave coda from a large earthquake require a network of broadband, high-dynamic-range seismic stations. Spectrograms (Peng et al., 2011a,b) together with visual scanning of high-pass-filtered data from such arrays greatly facilitate the identification of early triggered activity (see **Figure 4**; Prejean et al., 2004). This capability is unevenly distributed across the globe in existing seismic networks.
- *Statistical and physical significance of delayed triggering*. As the delay time between passing seismic waves and the onset of a suspected triggered response increases, it becomes increasingly difficult to establish statistical significance and a credible causal link between the two. Isolated earthquakes suspected of being triggered pose a particular problem in this regard regardless of delay times. Multiparameter monitoring networks capable of continuously tracking local changes in deformation, the hydrologic system, and other aseismic phenomena can provide important clues to aseismic processes that might link the perturbing dynamic stresses to a delayed or isolated seismic response – particularly when coupled with plausible physical models.
- *Distinguishing triggered earthquakes from aftershocks to triggered earthquakes*. The problem of distinguishing between static triggering and dynamic triggering of aftershocks to major earthquakes in the near-to-intermediate fields applies to remotely triggered earthquake swarms as well. In particular, deciding which earthquakes in a triggered swarm are directly triggered by teleseismic dynamic stresses and which earthquakes are aftershocks to the triggered earthquakes remains a challenge (Ziv, 2006). One promising approach to this problem involves the ETAS model described by Hainzl and Ogata (2005) to distinguish earthquakes driven by external forcing (such as fluid transport) from those resulting from an Omori-type aftershock sequence. High-resolution imaging of spatial–temporal patterns of swarm evolution can also provide evidence for fluid forcing, particularly when the seismicity front propagates as $(r/\sqrt{t})$ characteristic of fluid diffusion (Shapiro et al., 1997).
- *Quantitative physical models for the triggering process*. To determine the physical processes responsible for a remotely triggered earthquake, we need thorough knowledge of the triggering wave field in space and time together with precise hypocentral locations for the triggered seismicity and associated focal mechanisms. The more information available

on the physical environment under which triggering occurs, the better (e.g., fault orientations, frictional properties of faults, stress field, rock type, temperature, and the hydrologic regime). Strain transients have preceded or accompany episodes of triggered seismicity in volcanic regimes monitored by continuous, high-resolution strain data. Strain data are key in testing models that involve fluid activation and transport and whether triggered seismicity is a secondary response to a more fundamental aseismic process. High-frequency water well data will also be helpful in testing models that involve crustal fluid flow and pore pressure changes (Roeloffs, 1996), as these data can help determine how the crust responds hydrologically to large distant earthquakes.

4.11.7 Conclusions

Evidence in support of remote dynamic triggering as a common phenomenon is compelling. Remaining key questions include the following: what conditions favor dynamic triggering, and what physical processes do dynamic stresses induce in given crustal volumes that trigger local earthquakes and, at least in some cases, local aseismic deformation? To a large degree, emerging answers will come into focus as we obtain more complete sampling of the triggered response to future large earthquakes in terms of both spatial distribution and variations in response characteristics over a wide spectrum of timescales afforded by more widely distributed broadband seismic and continuous deformation instrumentation.

Based on the still uneven sampling available through mid-2013, it appears as though the seismogenic crust in tectonically active extensional regimes (such as the Basin and Range province in the western United States with strain rates of $\sim 10^{-8}$ year^{-1}) is more susceptible to dynamic triggering than stable cratonic regimes (such as the eastern United States with strain rates generally $< 10^{-9}$ year^{-1}). Higher tectonic strain rates mean shorter recovery (recharge) times between episodes of strain release (whether dynamically triggered or not). In volcanic and geothermal areas, recovery times may be further shortened by episodic advection of magmatic fluids into the crust from the lower crust or upper mantle that serve as local stress sources. Shorter recovery times mean that, at any given time, more areas are likely to be hovering in a near-critical state and thus susceptible to triggering by small dynamic stresses than in areas with low strain rates and extended recovery times. Elevated strain rates alone, however, are not sufficient. The seismogenic crust along the transpressional section of the San Andreas Fault through the California Coast Ranges with its high strain rate ($\sim 5 \times 10^{-7}$ year^{-1}) and frequent seismicity appears not to be particularly susceptible to dynamic triggering (Spudich et al., 1995), while the transition zone in the lower crust beneath the Parkfield section of the San Andreas Fault responds frequently with triggered tectonic tremor (Peng et al., 2009).

The transition zones at depths of 15–30 km along major plate-boundary megathrust and transform faults are particularly susceptible to frequent repeated dynamic triggering of tectonic tremor by seismic waves from large earthquakes and to modulation by tidal stresses. The triggered tremor occurs in the same volume as background tremor and ETS on megathrust faults. In both cases, the tremor appears to composed of closely spaced sequences of low-frequency, shear-failure earthquakes on extremely weak fault surfaces under near-lithostatic pore pressure and low effective friction coefficients < 0.1.

Models proposed to explain dynamic triggering fall into two broad categories: (1) frictional models and direct triggering of local seismicity and (2) models that appeal to the excitation of local crustal fluids or aseismic creep with seismicity developing as an indirect (and possibly delayed) response. For dynamic triggering under the frictional models, the stress state in the crust must differ from the Coulomb failure stress by less than the peak amplitudes of the dynamic stresses (typically < 10 kPa for remote triggering). These models are generally consistent with the rapid onset of triggered seismicity during the dynamic stresses followed by an Omori-like decay depending on the parameters for the specific friction law (e.g., rate-state and granular media).

Fluid excitation models involve fluid transport or a phase change driven by the triggered release of locally stored gravitational or chemical potential energy. These fluid-based models, whether hydrous or magmatic, together with models based on aseismic creep, involve some degree of crustal deformation through, for example, intrusion, pressure changes in the case of bubble excitation, advection in a magma body, or poroelasticity in the case of fluid diffusion. Triggered seismicity for many of these models can be a secondary response to quasi-static stresses generated by local deformation. These models admit the possibility of delayed onsets of triggered seismicity and increasing or sustained activity rates for extended periods following the passage of the dynamic stresses. In general, fluid-based models are likely to be more applicable in extensional stress regimes that host geothermal and volcanic systems than in stable regimes. Indeed, it is likely that frictional and fluid-driven processes both have some role in most instances of dynamic triggering.

Although currently available data are not yet adequate to clearly distinguish between the many competing models for the triggering process, the three-dimensional dynamic stress fields for seismic waves from large earthquakes or solid Earth tides, which can be calculated with reasonable accuracy, offer useful tools for investigating the state of stress in the crust and what it means to be in a near-critical state and providing insight on the earthquake nucleation process.

Acknowledgments

We are grateful to Tom Parsons and Zhigang Peng for constructive reviews of this manuscript and to Hiroo Kanamori for his guidance and patience as volume editor.

References

Aagaard BT, Anderson G, and Hudnut KW (2004) Dynamic rupture modeling of the transition from thrust to strike-slip motion in the 2002 Denali Fault earthquake, Alaska. *Bulletin of the Seismological Society of America* 94: S190–S201.

Aiken C, Peng Z, and Chao K (2013a) Tremors along the Queen Charlotte Margin triggered by large teleseismic earthquakes. *Geophysical Research Letters* 40(5): 829–834. http://dx.doi.org/10.1002/grl.50220.

Anderson JG, Brune JN, Louie JN, et al. (1994) Seismicity in the western Great Basin apparently triggered by the Landers, California, earthquake 28 June 1992. *Bulletin of the Seismological Society of America* 84: 863–891.

Anderson OL and Grew PC (1977) Stress corrosion theory of crack propagation with applications to geophysics. *Reviews of Geophysics* 15: 77–104.

Antonioli A, Cocco M, Das S, and Henery C (2002) Dynamic stress triggering during the great 25 March 1998 Antarctic Plate earthquake. *Bulletin of the Seismological Society of America* 92: 896–903.

Armstrong BH and Stierman DJ (1989) Acoustic emissions from foreshocks and secular strain changes prior to earthquakes. In: Reginald Hardy H Jr. (ed.) *Acoustic Emission/ Microseismic Activity in Geologic Structures and Materials*, vol. 17, pp. 309–326. Pennsylvania State University, State College PA: Trans Tech Publications.

Arnadottir T, Geirsson H, and Einarsson P (2004) Coseismic stress changes and crustal deformation on the Reykjanes Peninsula due to triggered earthquakes on 17 June 2000. *Journal of Geophysical Research* 109. http://dx.doi.org/10.1029/2004JB003130.

Atkinson BK (1984) Subcritical crack growth in geological materials. *Journal of Geophysical Research* 89: 4077–4114.

Bagdassarov N (1994) Pressure and volume changes in magmatic systems due to the vertical displacement of compressible materials. *Journal of Volcanology and Geothermal Research* 63: 95–100.

Bakun WH and McGarr A (2002) Differences in attenuation among stable continental regions. *Geophysical Research Letters* 29: 36-1–36-4.

Beeler NM, Simpson RW, Hickman SH, and Lockner DA (2000) Pore fluid pressure, apparent friction and Coulomb failure. *Journal of Geophysical Research* 105(B11): 25533–25542.

Beeler NM and Lockner DA (2003) Why earthquakes correlate weakly with the solid Earth tides: Effects of periodic stress on the rate and probability of earthquake occurrence. *Journal of Geophysical Research* 108. http://dx.doi.org/10.1029/2001JB001518.

Beresnev IA and Wen K-L (1995) Remotely triggered seismicity inferred from Taiwan regional catalog. *Geophysical Research Letters* 22: 3155–3158.

Beroza GC and Ide S (2011) Slow earthquakes and nonvolcanic tremor. *Annual Review of Earth and Planetary Sciences* 39: 271–296.

Brodsky EE (2006) Long-range triggered earthquakes that continue after the wave train passes. *Geophysical Research Letters* 33. http://dx.doi.org/10.1029/2006GL026605.

Brodsky EE, Karakostas V, and Kanamori H (2000) A new observation of dynamically triggered regional seismicity: Earthquakes in Greece following the August, 1999, Izmit, Turkey earthquake. *Geophysical Research Letters* 27: 2741–2744.

Brodsky EE and Prejean SG (2005) New constraints on mechanisms of remotely triggered seismicity at Long Valley Caldera. *Journal of Geophysical Research* 110. http://dx.doi.org/10.1029/2004JB003211.

Brodsky EE, Roeloffs E, Woodcock D, Gall I, and Manga M (2003) A mechanism for sustained groundwater pressure changes induced by distant earthquakes. *Journal of Geophysical Research* 108. http://dx.doi.org/10.1029/2002JB002321.

Brodsky E, Sturtevant B, and Kanamori H (1998) Earthquakes, volcanoes, and rectified diffusion. *Journal of Geophysical Research* 103: 23827–23838.

Byerlee JD (1980) Friction of rocks. *Pure and Applied Geophysics* 116: 615–626.

Camelbeeck T, van Eck T, Pelzing R, et al. (1994) The 1992 Roermond earthquake, the Netherlands, and its aftershocks. *Geologie en Mijnbouw* 73: 181–197.

Cannata A, Di Grazia G, Montalto P, Aliota M, Patane D, and Boschi E (2010) Response of Mount Etna o dynamic stresses from distant earthquakes. *Journal of Geophysical Research* 115: B12304. http://dx.doi.org/10.1029/2010JB007487.

Chao K, Peng Z, Fabian A, and Ojha L (2012) Comparisons of triggered tremor in California. *Bulletin of the Seismological Society of America* 102(2): 900–908.

Chao K, Peng Z, Gonzalez-Huizar H, et al. (2013) Global search of triggered tremor following the 2011 Mw9.0 Tohoku-Oki earthquake. *Bulletin of the Seismological Society of America* 103(2b). http://dx.doi.org/10.1785/0130120171.

Chao K, Peng Z, Wu C, Tang C-C, and Lin C-H (2011) Remote triggering of non-volcanic tremor around Taiwan. *Geophysical Journal International* 188: 301–324.

Chesley C, LaFemina PC, Puskas C, and Kobayashi D (2012) The 1707 Mw8.7 Hoei earthquake triggered the largest historical eruption of Mt. Fuji. *Geophysical Research Letters* 39: L224309.

Chouet B (1992) A seismic model for the source of long-period events and harmonic tremor. In: Gasparini P, Scarpa R, and Aki K (eds.) *Volcanic Seismology*pp. 133–156. Berlin: Springer-Verlag.

Chouet B, Dawson P, and Nakano M (2006) Dynamics of diffusive bubble growth and pressure recovery in a bubbly rhyolitic melt embedded in an elastic solid. *Journal of Geophysical Research* 111: B07310. http://dx.doi.org/10.1029/2005JB004174.

Christiansen L, Hurwitz S, Saar MO, Ingebritsen SE, and Hsieh P (2005) Seasonal seismicity at western United States volcanic centers. *Earth and Planetary Science Letters* 240: 307–321.

Cocco M and Rice JR (2002) Pore pressure and poroelasticity effects in Coulomb stress analysis of earthquake interactions. *Journal of Geophysical Research* 107. http://dx.doi.org/10.1029/2002JB002319.

Cochran ES, Vidale JE, and Tanaka S (2004) Earth tides can trigger shallow thrust fault earthquakes. *Science* 306: 1164–1166.

Daniel D, Marsan D, and Bouchon M (2008) Earthquake triggering in southern Iceland following the June 2000 M_S 6.6 doublet. *Journal of Geophysical Research* 113: B05310. http://dx.doi.org/10.1029/2007JB005107.

Das S and Scholz C (1981) Off-fault aftershock clusters caused by shear stress increase? *Journal of Geophysical Research* 71: 1669–1675.

Daub EG, Shelly DR, Guyer RA, and Johnson PA (2011) Brittle and ductile friction and the physics of tectonic tremor. *Geophysical Research Letters* 38: L10301. http://dx.doi.org/10.1029/GL046866.

Delle-Donne D, Harris AJL, Ripepe M, and Wright R (2010) Earthquake-induced thermal anomalies at active volcanoes. *Geology* 38: 771–774.

Di Carli S, Voisin C, Cotton F, and Semmane F (2008) The 2000 western Tottori (Japan) earthquake: Triggering of the largest aftershock and constraints on the slip-weakening distance. *Journal of Geophysical Research* 113: B05307. http://dx.doi.org/10.1029/2007JB004951.

Dieterich JH (1979) Modeling of rock friction 1. Experimental results and constitutive equations. *Journal of Geophysical Research* 84: 2161–2168.

Durand V, Bouchon M, Karabulut H, Marsan D, and Schmittbuhl J (2013) Link between Coulomb stress changes and seismic activation in the eastern Marmara sea after the 1999, Izmit (Turkey), earthquake. *Journal of Geophysical Research: Solid Earth* 118: 681–688.

Durand V, Bouchon M, Karabulut H, et al. (2010) Seismic interaction and delayed triggering along the north Anatolian Fault. *Geophysical Research Letters* 37: L18310. http://dx.doi.org/10.1029/201GL044688.

Eberhart-Phillips D, Haeussler PJ, Freymueller JT, et al. (2003) The 2002 Denali fault earthquake, Alaska: A large magnitude, slip-partitioned event. *Nature* 300: 1113–1118.

Elkhoury JE, Niemeijer A, Brodsky E, and Marone C (2011) Laboratory observations of permeability enhancement by fluid pressure oscillations of an in situ fractured rock. *Journal of Geophysical Research* 116: B02311. http://dx.doi.org/10.1029/2010BJ007759.

Emiliani C, Harrison CGA, and Swanson M (1969) Underground nuclear explosions and the control of earthquakes. *Science* 165: 1255–1256.

Enescu B, Chao K, Peng Z, et al. (2012) Love wave triggering of non-volcanic tremor in the Nankai region, southwest Japan: Observations and physical interpretation. In: *Abstract S33B-2550 presented at 2012 Fall Meeting, 3–7 December, AGU, San Francisco, CA* .

Felzer KR and Brodsky EE (2005) Testing the stress shadow hypothesis. *Journal of Geophysical Research* 110. http://dx.doi.org/10.1029/2004JB003277.

Felzer KR and Brodsky EE (2006) Evidence for dynamic aftershock triggering from earthquake densities. *Nature* 441: 735–738.

Fournier RO (1999) Hydrothermal processes related to movement of fluid from plastic to brittle rock in the magmatic-epithermal environment. *Economic Geology* 94: 1193–1211.

Freed AM (2005) Earthquake triggering by static, dynamic, and postseismic stress transfer. *Annual Review of Earth and Planetary Sciences* 33: 335–367.

Fry B, Chao K, Bannister S, and Peng Z (2011) Deep tremor beneath the Hikurangi margin in New Zealand triggered by the 2010 M_w 8.8 Chile earthquake. *Geophysical Research Letters* 38: L15306.

Galperin EI, Petersen NV, Sitnikov AV, and Vinnik LP (1990) On the properties of short-period seismic noise. *Physics of the Earth and Planetary Interiors* 63: 163–171.

Ghosh A, Vidale JE, Peng Z, Creager KC, and Houston H (2009) Complex nonvolcanic tremor near Parkfield, California, triggered by the great 2004 Sumatra earthquake. *Journal of Geophysical Research* 114. http://dx.doi.org/10.1029/2008JB006062.

Glowacka E, Nava AF, Cossio DD, Wong V, and Farfan F (2002) Fault slip, seismicity, and deformation in the Mexicali Valley, Baja California, Mexico, after the M 7.1 Hector Mine earthquake. *Bulletin of the Seismological Society of America* 92: 1290–1299.

Gomberg J (1996) Stress/strain changes and triggered seismicity following the Mw 7.3 Landers, California, earthquake. *Journal of Geophysical Research* 101: 751–764.

Gomberg J (2010) Lessons from (triggered) tremor. *Journal of Geophysical Research* 115: B10302. http://dx.doi.org/10.1029/2009JB007011.

Gomberg J, Blanpied ML, and Beeler NM (1997) Transient triggering of near and distant earthquakes. *Bulletin of the Seismological Society of America* 87: 294–309.

Gomberg J, Bodin P, Larson K, and Dragert H (2004) Earthquakes nucleated by transient deformations caused by the M = 7.9 Denali, Alaska, earthquake. *Nature* 427: 621–624.
Gomberg J, Bodin P, and Reasenberg PA (2003) Observing earthquakes triggered in the near field by dynamic deformations. *Bulletin of the Seismological Society of America* 93: 118–138.
Gomberg J and Davis S (1996) Stress/strain changes and triggered seismicity at the Geysers, California. *Journal of Geophysical Research* 101: 733–749.
Gomberg J and Johnson P (2005) Dynamic triggering of earthquakes. *Nature* 437: 830.
Gomberg J, Reasenberg PA, Bodin P, and Harris R (2001) Earthquakes triggering by seismic waves following the Landers and Hector mine earthquakes. *Nature* 411: 462–465.
Gomberg J, Reasenberg PA, Cocco M, and Belardinelli ME (2005) A frictional population model of seismicity rate change. *Journal of Geophysical Research* 110. http://dx.doi.org/10.1029/2004JB003404.
Gomberg J, Rubinstein JL, Peng Z, Crager KC, Vidale JE, and Bodin P (2008) Widespread triggering of nonvolcanic tremor in California. *Science* 319: 173.
Gomberg J and Sherrod B (2014) Crustal earthquake triggering by modern great earthquakes on subduction zone thrusts. *Journal of Geophysical Research: Solid Earth* 119.2: 1235–1250.
Gonzalez-Huizar H and Velasco AA (2011) Dynamic triggering: Stress modeling and a case study. *Journal of Geophysical Research* 116. http://dx.doi.org/10.10/2009JB007000.
Gonzalez-Huizar H, Velasco AA, Peng Z, and Castro RR (2012) Remote triggered seismicity caused by the 2011, M9.0 Tohoku-Oki, Japan, earthquake. *Geophysical Research Letters* 39: L10302.
Guillhem A, Peng Z, and Nadeau RM (2010) High-frequency identification of non-volcanic tremor triggered by regional earthquakes. *Geophysical Research Letters* 37. http://dx.doi.org/10.10/2010GL044660.
Guillhem A and Nadeau RM (2012) Episodic tremor and deep slow slip events in central California. *Earth and Planetary Science Letters* 357: 1–10.
Hainzl S and Ogata Y (2005) Detecting fluid signals in seismicity data through statistical earthquake modeling. *Journal of Geophysical Research* 110. http://dx.doi.org/10.1029/2004JB003247.
Harrington RM and Brodsky EE (2006) The absence of remotely triggered seismicity in Japan. *Bulletin of the Seismological Society of America* 96: 871–878.
Harris RA (1998) Introduction to a special section: Stress triggers, stress shadows, and implications for seismic hazards. *Journal of Geophysical Research* 103: 24347–24358.
Harris RA and Day SM (1993) Dynamics of fault interaction: Parallel strike-slip faults. *Journal of Geophysical Research* 98: 4461–4472.
Harris RA, Dolan JF, Hartlib R, and Day SM (2002) The 1999 Izmit, Turkey, earthquake: A 3-D dynamic stress transfer model of interearthquake triggering. *Bulletin of the Seismological Society of America* 92: 245–255.
Hawthorne JC and Rubin AM (2010) Tidal modulation of slow slip in Canada. *Journal of Geophysical Research* 115: B09406.
Hill DP (2008) Unrest in Long Valley Caldera, California: 1978–2004. In: De Natale G, Troise C, and Kilburn CRJ (eds.) *Mechanisms of Activity and Unrest at Large Calderas*, pp. 1–24. London: Geological Society. Publication No. 269.
Hill DP (2012a) Dynamic stresses Coulomb failure, and remote triggering – Corrected. *Bulletin of the Seismological Society of America* 102. http://dx.doi.org/10.1785/0120120085.
Hill DP (2012b) Surface-wave potential for triggering tectonic (nonvolcanic) tremor – Corrected. *Bulletin of the Seismological Society of America* 102. http://dx.doi.org/10.1785/0120120086.
Hill DP, Johnston MJS, Langbein JO, and Bilham R (1995) Response of Long Valley Caldera to the Mw = 7.3 Landers, California, earthquake. *Journal of Geophysical Research* 100: 12985–13005.
Hill DP, Peng Z, Shelly DR, and Aiken C (2013) S-wave triggering of tremor beneath the Parkfield, CA, section of the San Andreas Fault by the 2011 Tohoku-Oki Japan earthquake: Observations and theory. *Bulletin of the Seismological Society of America* 103. http://dx.doi.org/10.1785/0120120114.
Hill DP, Pollitz F, and Newhall C (2002) Earthquake-volcano interactions. *Physics Today* 55: 41–47.
Hill DP and Prejean SG (2007) Dynamic triggering. In: Kanamori H (ed.) *Treatise on Geophysics. Earthquake Seismology*, vol. 4, pp. 257–291. Amsterdam: Elsevier.
Hill DP, Reasenberg PA, Michael A, et al. (1993) Seismicity remotely triggered by the magnitude 7.3 Landers, California, earthquake. *Science* 260: 1617–1623.
Hirose F, Miyaoka K, Hayashimoto N, Yamazaki T, and Nakamura M (2011) Outline of the 2011 off the Pacific coast of Tohoku Earthquake (Mw 9.0) – seismicity: Foreshocks, mainshock, aftershocks, and induced activity. *Earth, Planets and Space* 63: 513–518.
Hough SE (2001) Triggered earthquakes and the 1811–1812 New Madrid, central United States, earthquake sequence. *Bulletin of the Seismological Society of America* 91: 1547–1581.
Hough SE (2005) Remotely triggered earthquakes following moderate mainshocks (or why California is not falling into the ocean). *Seismological Research Letters* 76: 58–66.
Hough SE, Billham R, Ambraseys N, and Feldl N (2005) Revisiting the 1897 Shillong and 1905 Kangra earthquakes in northern India: Site response, Moho reflections and a triggered earthquake. *Current Science* 88: 1632–1638.
Hough SE and Kanamori H (2002) Source properties of earthquakes near the Salton Sea triggered by the 16 October 1999 Mw 7.1 Hector Mine, California, earthquake. *Bulletin of the Seismological Society of America* 92: 1281–1289.
Hough SE, Seeber L, and Armbruster JG (2003) Intraplate triggered earthquakes: Observations and interpretation. *Bulletin of the Seismological Society of America* 93: 2212–2221.
Hudnut KWL, Seeber L, and Pacheco J (1989) Cross-fault triggering in the November 1978 Superstition Hill earthquake sequence, southern California. *Geophysical Research Letters* 16: 199–202.
Husen S, Taylor R, Smith RB, and Healser H (2004a) Changes in geyser eruption behavior and remotely triggered seismicity in Yellowstone National Park produced by the 2002 M 7.9 Denali fault earthquake, Alaska. *Geology* 32: 537–540.
Husen S, Wiemer S, and Smith RB (2004b) Remotely triggered seismicity in the Yellowstone National Park region by the 2002 Mw 7.9 Denali Fault earthquake, Alaska. *Bulletin of the Seismological Society of America* 94: S317–S331.
Husker AL and Brodsky EE (2004) Seismicity in Idaho and Montana triggered by the Denali Fault earthquake: A window into the geologic context for seismic triggering. *Bulletin of the Seismological Society of America* 94: S310–S316.
Ichihara M and Brodsky EE (2006) A limitation on the effect of rectified diffusion in volcanic systems. *Geophysical Research Letters* 33. http://dx.doi.org/10.1029/2005GL024733.
Ide S (2010) Striations, duration, migration and tidal response in deep tremor. *Nature* 466: 356–360.
Jay JA, Pritchard ME, West ME, et al. (2012) Shallow seismicity, triggered seismicity, and ambient noise tomography at the long-dormant Uturuncu volcano. *Bulletin of Volcanology* 74: 817–837. http://dx.doi.org/10.1007/s00445-011-0568-7.
Jiang T, Peng Z, Wang W, and Chen Q-F (2010) Remotely triggered seismicity in continental China following the 2008 Mw 7.9 Wenchuan earthquake. *Bulletin of the Seismological Society of America* 100: 2574–2589.
Johnson PA, Carpenter BM, Knuth M, et al. (2012) Nonlinear dynamical triggering of slow-slip on simulated earthquake faults. *Journal of Geophysical Research* 117: B04310. http://dx.doi.org/10.1029/2011JB 008594.
Johnson PA and Jia X (2005) Nonlinear dynamic, granular media and dynamic earthquake triggering. *Nature* 437: 871–874.
Johnson PA, Savage M, Knuth K, and Marone C (2008) Effects of acoustic waves on stick–slip in granular media and implications for earthquakes. *Nature* 451: 57–60.
Johnston MJS, Hill DP, Linde AT, Langbein J, and Bilham R (1995) Transient deformation during triggered seismicity from the 28 June 1992 Mw = 7.3 Landers earthquake at Long Valley volcanic caldera, California. *Bulletin of the Seismological Society of America* 85: 787–795.
Johnston MJS, Hill DP, and Pitt AM (2000) Strain transient recorded in the Long Valley Caldera during triggered seismicity from the October 16, 1999 M7.1 Hector Mine, California, earthquake. *EOS, Transactions of the American Geophysical Union* 81, WP1384.
Johnston MJS, Prejean SG, and Hill DP (2004a) Triggered deformation and seismic activity under Mammoth Mountain in Long Valley caldera by the 3 November 2002 Mw 7.9 Denali Fault earthquake. *Bulletin of the Seismological Society of America* 94: S360–S369.
Johnston MJS, Prejean S, and Hill DP (2004b) Triggered deformation and seismic activity in Long Valley Caldera by the November 3, 2002, M 7.9 Denali Fault earthquake. *Bulletin of the Seismological Society of America* 94: S360–S369.
Jousset P and Rohmer J (2012) Evidence for remotely triggered microearthquakes during salt cavern collapse. *Geophysical Journal International* 191: 207–223.
Julian BR (1994) Volcanic tremor: Nonlinear excitation by fluid flow. *Journal of Geophysical Research* 99: 11859–11877.
Julian BR (2002) Seismological detection of slab metamorphism. *Science* 296: 1625–1626.
Kanamori H and Brodsky EE (2004) The physics of earthquakes. *Reports in Progress in Physics* 67: 1429–1496.
Kilb D, Gomberg J, and Bodin P (2000) Triggering of earthquake aftershocks by dynamic stresses. *Nature* 408: 570–574.
Kilb D, Gomberg J, and Bodin P (2002) Aftershock triggering by complete Coulomb stress changes. *Journal of Geophysical Research* 107. http://dx.doi.org/10.1029/2001JB000202.

King GCP and Cocco M (2001) Fault interactions by elastic stress changes: New clues from earthquake sequences. *Advances in Geophysics* 44: 1–38.

Kisslinger C and Jones LM (1991) Properties of aftershock sequences in southern California. *Journal of Geophysical Research* 96: 11947–11958.

Klein FW (1976) Earthquake swarms and the semidiurnal solid earth tide. *Geophysical Journal of the Royal Astronomical Society* 54: 245–295.

Lapusta N and Rice JR (2003) Nucleation and early seismic propagation of small and large events in a crustal earthquake model. *Journal of Geophysical Research* 108. http://dx.doi.org/10.1029/2001JB000793.

Lay T, Kanamori H, Ammon CJ, et al. (2005) The great Sumatra-Andaman earthquake of 26 December 2004. *Science* 308: 1127–1132.

Lei X, Xie C, and Fu B (2011) Remotely triggered seismicity in Yunnan, southwestern China, following the 2004 M_w 9.3 Sumatra earthquake. *Journal of Geophysical Research* 116: B08303. http://dx.doi.org/10.1029/2011JB008245.

Lin C-H (2012) Remote triggering of the Mw 6.9 Hokkaido earthquake as a result of the Mw 6.6 Indonesian Earthquake on September 11, 2008. *Terrestrial, Atmospheric and Oceanic Sciences* 23: 283–290.

Linde AT and Sacks IS (1998) Triggering of volcanic eruptions. *Nature* 395: 888–890.

Linde AT, Sacks IS, Johnston MJS, Hill DP, and Billham RG (1994) Increased pressure from rising bubbles as a mechanism for remotely triggered seismicity. *Nature* 371: 408–410.

Lockner DA and Beeler NM (1999) Premonitory slip and tidal triggering of earthquakes. *Journal of Geophysical Research* 104: 20133–20151.

Lockner DA and Beeler NM (2002) Rock failure and earthquakes. In: Lee WHK, Kanamori H, Jennings PC, and Kisslinger C (eds.) *International Handbook of Earthquake and Engineering Seismology Part A*, pp. 505–537. Amsterdam: Academic Press.

Luttrell K and Sandwell D (2010) Ocean loading effects on stress at near shore boundary fault systems. *Journal of Geophysical Research* 115: B08411. http://dx.doi.org/10.1029/2009JB006541.

Manga M and Brodsky EE (2006) Seismic triggering of eruptions in the far field: Volcanoes and geysers. *Annual Review of Earth and Planetary Sciences* 34: 263–291.

Manga M and Wang C-Y (2007) Earthquake hydrology. In: Kanamori H (ed.) *Treatise on Geophysics. Earthquake Seismology*, vol. 4, pp. 293–320. Amsterdam: Elsevier.

Mangan M and Sisson T (2000) Delayed, disequilibrium degassing in rhyolite magma: Decompression experiments and implications for explosive volcanism. *Earth and Planetary Science Letters* 183: 441–455.

Marsan D (2003) Triggering of seismicity at short distances following California earthquakes. *Journal of Geophysical Research* 108. http://dx.doi.org/10.1029/2002JB001946.

Marzocchi W (2002) Remote seismic influence on large explosive eruptions. *Journal of Geophysical Research* 107. http://dx.doi.org/10.1029/2001JB000307.

Matthews MV and Reasenberg PA (1988) Statistical methods for investigating quiescence and other temporal seismicity patterns. *Pure and Applied Geophysics* 126: 357–372.

McGarr A, Simpson D, and Seeber L (2002) Case histories of induced and triggered seismicity. In: Lee WHK, Kanamori H, Jennings PC, and Kisslinger C (eds.) *International Handbook of Earthquake and Engineering Seismology, Part A*, pp. 647–664. Amsterdam: Academic Press.

Meltzner AJ and Wald DJ (2003) Aftershocks and triggered events of the great 1906 California earthquake. *Bulletin of the Seismological Society of America* 93: 2160–2186.

Meng X, Peng Z, and Hardebeck J (2013a) Seismicity around Parkfield correlates with static shear stress changes following the 2003 Mw6.5 San Simeon earthquake. *Journal of Geophysical Research* 118: 3576–3591. http://dx.doi.org/10.1002/jgrb.50271.

Meng X, Peng Z, Withers K, Olsen K, Yu X, and Hong B (2013b) Systematic search of missing earthquakes near the Salton Sea geothermal field and San Jacinto fault around the 2010 Mw7.2 El Mayor-Cucapah earthquake. (abs.) *Seismological Research Letters* 84(2): 391.

Miller SA (2008) Note on rain-triggered earthquakes and their dependence on karst geology. *Geophysical Journal International* 173: 334–338.

Miyazawa M (2011) Propagation of an earthquake triggering front from the 2011 Tohoku-oki earthquake. *Geophysical Research Letters* 38: L23307. http://dx.doi.org/10.1029/2011GL049795.

Miyazawa M (2012) Detection of seismic events triggered by P-waves from the 2011 Tohoku-oki earthquake. *Earth, Planets and Space* 64: 1223–1229.

Miyazawa M and Brodsky E (2008) Deep low-frequency tremor that correlates with passing surface waves. *Journal of Geophysical Research* 113. http://dx.doi.org/10.1029/2006JB004890.

Miyazawa M, Brodsky E, and Mori J (2008) Learning from dynamic triggering of low-frequency tremor in subduction zones. *Earth, Planets and Space* 60: e17–e20.

Miyazawa M and Mori J (2005) Detection of triggered deep low-frequency events from the 2003 Tokachi-oki earthquake. *Geophysical Research Letters* 32: L10307.

Miyazawa M and Mori J (2006) Evidence suggesting fluid flow beneath Japan due to periodic seismic triggering from the 2004 Sumatra-Andaman earthquake. *Geophysical Research Letters* 33: L05303-L. http://dx.doi.org/10.1029/2005GL025087.

Miyazawa M, Nakanishi I, Sudo Y, and Ohkura T (2005) Dynamic response of frequent tremors at Aso volcano to teleseismic waves from the 1999 Chi-Chi, Taiwan earthquake. *Journal of Volcanology and Geothermal Research* 147: 173–186.

Mohamad R, Darkal AN, Seber D, Sandoval E, Gomez F, and Barazangi M (2000) Remote earthquake triggering along the Dead Sea Fault in Syria following the 1995 Gulf of Aqaba earthquake (Ms = 7.3). *Seismological Research Letters* 71: 47–52.

Moran SC (2003) Multiple seismogenic processes for high-frequency earthquakes at Katmai National Park, Alaska: Evidence from stress tensor inversions of fault plane solutions. *Bulletin of the Seismological Society of America* 93: 94–108.

Moran SC, Power JA, Stihler SD, Sanchez JJ, and Caplin-Auerback J (2004) Earthquake triggering at Alaskan volcanoes following the 3 November 2002 Denali Fault earthquake. *Bulletin of the Seismological Society of America* 94: S300–S309.

Moran SC, Zimbelman DR, and Malone SD (2000) A model for the magmatic-hydrothermal system at Mount Rainier, Washington, from seismic and geochemical observations. *Bulletin of Volcanology* 61: 425–436.

Mori J and Kano Y (2009) Is the 2006 Yogyakarta earthquake related to the triggering of the Sidoarjo, Indonesia mud volcano? *Journal of Geography* 118: 492–498.

Mori J, McGee C, Talai B, and Itikarai I (1989) A summary of precursors to volcanic eruptions in Papua New Guinea. In: Latter JH (ed.) *Volcanic Hazards Assessment and Monitoring*, pp. 260–291. Berlin: Springer-Verlag.

Nakata R, Suda N, and Tsuruoka H (2008) Non-volcanic tremor resulting from the combined effect of Earth tides and slow slip events. *Nature Geoscience* 1: 676–678.

Nicolaev AV and Troitskii PA (1987) Lithospheric studies based on array analysis of P-waves and microseisms. *Tectonophysics* 140: 103–113.

Nostro C, Stein RS, Cocco M, Belardinelli ME, and Marzocchi W (1998) Two-way coupling between Vesuvius eruptions and southern Apennine earthquakes, Italy, by elastic stress transfer. *Journal of Geophysical Research* 103(B10): 24487–24504.

Nur A and Booker JR (1972) Aftershocks caused by pore fluid flow? *Science* 175: 885–887.

Obara K (2002) Nonvolcanic deep tremor associated with subduction in southwest Japan. *Science* 296: 1679–1681.

Obara K (2003) Time sequence of deep low-frequency tremors in the southwest Japan subduction zone: Triggering phenomena and periodic activity (in Japanese). *Journal of Geography (Japan)* 112: 837–849.

Omori F (1894) On the aftershocks of earthquakes. *Journal of the College of Science, Imperial University of Tokyo* 7: 111–120.

Oppenheimer DH, Reasenberg PA, and Simpson RW (1988) Fault plane solutions for the 1984 Morgan Hill, California, earthquake sequence: Evidence for the state of stress on the Calaveras fault. *Journal of Geophysical Research* 93: 9007–9026.

Pankow KL, Arabasz WJ, Pechmann JC, and Nava SJ (2004) Triggered seismicity in Utah from the 3 November 2002 Denali fault earthquake. *Bulletin of the Seismological Society of America* 94: S332–S347.

Parsons T (2005) A hypothesis for delayed dynamic earthquake triggering. *Geophysical Research Letters* 32: L04302. http://dx.doi.org/10.1029/2004GL021811.

Parsons T, Kaven JO, Velasco AA, and Gonzalez-Huizar H (2012) Unraveling the apparent magnitude threshold of remote earthquake triggering using full wave-field surface wave simulations. *Geochemistry, Geophysics, Geosystems* 13: Q06016. http://dx.doi.org/10.1029/2012GC004164.

Parsons T and Velasco AA (2009) On near-source earthquake triggering. *Journal of Geophysical Research* 114: B10307. http://dx.doi.org/10.1029/2008JB006277.

Parsons T and Velasco AA (2011) Absence of remotely triggered large earthquakes beyond the mainshock region. *Nature Geoscience* 4: 312–316.

Peng Z and Chao K (2008) Non-volcanic tremor beneath the Central Range in Taiwan triggered by the 2001 Mw 7.8 Kunlun earthquake. *Geophysical Journal International* 175: 825–829.

Peng Z, Chao K, Wu C, Fry B, Enescu B, and Aiken C (2012) Global observations of triggered tectonic tremor. *Seismological Research Letters* 83(2): 417.

Peng Z, Gonzalez-Huizar H, Chao K, Aiken C, Moreno B, and Armstrong G (2013) Tectonic tremor beneath Cuba triggered by the M_W 8.8 Maule and M_W 9.0 Tohoku-Oki earthquakes. *Bulletin of the Seismological Society of America* 103(1): 595–600.

Peng Z, Hill DP, Shelly DR, and Aiken C (2010a) Remotely triggered microearthquakes and tremor in central California following the 2010 Mw 8.8 Chile earthquake. *Geophysical Research Letters* 37: L24314.

Peng Z, Long LT, and Zhao P (2011a) The relevance of high-frequency analysis artifacts to remote triggering. *Seismological Research Letters* 82(5): 662–856. http://dx.doi.org/10.1785/gssrl.82.5.656.

Peng Z, Vidale JE, Creager KC, Rubinstein JL, Gomberg J, and Bodin P (2008) Strong tremor near Parkfield, CA, excited by the 2002 Denali fault earthquake. *Geophysical Research Letters* 35: L23305. http://dx.doi.org/10.1029/2008GL036080.

Peng Z, Vidale JE, Wech AG, Nadeau RM, and Creager KC (2009) Remote triggering of tremor along the San Andreas fault in central California. *Journal of Geophysical Research* 114. http://dx.doi.org/10.1029/2008JB006049.

Peng Z, Walter J, Aster R, Nyblade A, Wiens D, and Anandakrishnan S (2014) Antarctic icequakes triggered by the 2010 Maule earthquake in Chile. *Nature Geoscience* (in press).

Peng Z, Wang W, Chen Q-F, and Jiang T (2010b) Remotely triggered seismicity in northeast China following the 2008 Mw 7.9 Wenchuan earthquake. *Earth, Planets and Space* 62: 893–898.

Peng Z, Wu C, and Aiken C (2011b) Delayed triggering of microearthquakes by multiple surface waves circling the Earth. *Geophysical Research Letters* 38: L04306. http://dx.doi.org/10.1029/2010GL046373.

Perfettini HJ, Schmittbuhl J, and Cochard A (2003) Shear and normal load perturbations on a two-dimensional continuous fault: 2. Dynamic triggering. *Journal of Geophysical Research* 108.

Pollitz FF and Johnston MJS (2006) Direct test of static stress versus dynamic stress triggering of aftershocks. *Geophysical Research Letters* 33. http://dx.doi.org/10.1029/2006GL026764.

Pollitz FF and Sacks IS (2002) Stress triggering of the 1999 Hector Mine earthquake by transient deformation following the 1992 Landers earthquake. *Bulletin Seismological Society of America* 92: 1487–1496.

Pollitz FF, Stein RS, Sevilgen V, and Burgmann R (2012) The 11 April 2012 east Indian Ocean earthquake triggered large aftershocks worldwide. *Nature* 490: 250–253.

Power JA, Moran SC, McNutt SR, Stihler SD, and Sanchez JJ (2001) Seismic response of the Katmai volcanoes to the 6 December 1999 magnitude 7.0 Karluk Lake earthquake, Alaska. *Bulletin of the Seismological Society of America* 91: 57–63.

Prejean SG and Hill DP (2009) Earthquakes, dynamic triggering of. In: Meyers R (ed.) *Complexity in Earthquakes, Tsunamis, and Volcanoes, and Forecast*, pp. 2600–2621. Berlin: Springer.

Prejean SG, Hill DP, Brodsky EE, et al. (2004) Remotely triggered seismicity on the United States west coast following the Mw 7.9 Denali fault earthquake. *Bulletin of the Seismological Society of America* 94: S348–S359.

Pyle DM and Pyle DL (1995) Bubble migration and the initiation of volcanic eruptions. *Journal of Volcanology and Geothermal Research* 67: 227–232.

Richards-Dinger K, Stein RS, and Toda S (2010) Decay of aftershock density with distance does not indicate triggering by dynamic stress. *Nature* 467: 583–588.

Richter CF (1958) *Elementary Seismology*. San Francisco, CA: Freeman and Co, 768 p.

Rivera L and Kanamori H (2002) Spatial heterogeneity of tectonic stress in the crust. *Geophysical Research Letters* 39. http://dx.doi.org/10.1029/2001GL013803.

Roeloffs E (1996) Poroelastic techniques in the study of earthquake-related hydrologic phenomena. *Advances in Geophysics* 37: 135–195.

Rogers D and Draggert H (2003) Episodic tremor and slip on the Cascadia subduction zone: The chatter of slip. *Science* 300: 1942–1943.

Rowe C, Christensen D, and Carver GA (2004) Preface to the issue dedicated to the 2002 Denali fault earthquake sequence. *Bulletin of the Seismological Society of America* 94: S1–S4.

Rubinstein JL, Gomberg J, Vidale JE, et al. (2009) Seismic wave triggering of nonvolcanic tremor, episodic tremor and slip, and earthquakes on Vancouver Island. *Journal of Geophysical Research* 114: B00A01. http://dx.doi.org/10.1029/2008JB005875.

Rubinstein JL, La Rocca M, Vidale JE, Creager KC, and Wech AG (2008) Tidal modulation of non-volcanic tremor. *Science* 319: 186–189.

Rubinstein JL, Shelly DR, and Ellsworth WL (2010) Non-volcanic tremor: A window into the roots of fault zones. In: Cloetingh D and Negendank J (eds.) *New Frontiers in Integrated Solid Earth Science*, pp. 287–314. Springer: Dordrecht.

Rubinstein JL, Vidale JE, Gomberg J, Bodin P, Creager KC, and Malone SD (2007) Non-volcanic tremor driven by large transient shear stresses. *Nature* 448: 579–582.

Rybicki K, Kato T, and Kasahara K (1985) Mechanical interaction between neighboring active faults – static and dynamic stress field induced by faulting. *Bulletin of the Earthquake Research Institute, University of Tokyo* 60: 1–21.

Sahagian DL and Proussevitch AA (1992) Bubbles in volcanic systems. *Nature* 359: 485.

Sanchez JJ and McNutt SR (2004) Intermediate-term declines in seismicity at Mt. Wrangell and Mt. Veniaminof volcanoes, Alaska, following the 3 November 2002 Mw 7.9 Denali Fault earthquake. *Bulletin of the Seismological Society of America* 94: S370–S383.

Savage JC and Clark MM (1982) Magmatic resurgence in Long Valley Caldera, eastern California. *Science* 217: 531–533.

Savage HM and Morone C (2008) Potential for earthquake triggering from transient deformations. *Journal of Geophysical Research* 113: B05302. http://dx.doi.org/10.1029/2007JB005277.

Scholz CH (1998) Earthquakes and friction laws. *Nature* 391: 37–42.

Shapiro SA, Huenges E, and Borm G (1997) Estimating the crust permeability from fluid-injection-induced seismic emission at the KTB site. *Geophysical Journal International* 131: F15–F18.

Shelly DR, Beroza GC, and Ide S (2007) Complex evolution of transient slip derived from precise tremor locations in western Shikoku, Japan. *Geochemistry, Geophysics, Geosystems* 8: Q100014. http://dx.doi.org/10.1029/2007GC001640.

Shelly DR, Beroza GC, Ide S, and Nakamula S (2006) Low-frequency earthquakes and their relation to episodic tremor and slip. *Nature* 442: 188–191.

Shelly DR, Moran SC, and Thelen WA (2013) Evidence for fluid-triggered slip in the 2009 Mount Rainier, Washington, earthquake swarm. *Geophysical Research Letters* 40. http://dx.doi.org/10.1002/grl.50354.

Shelly DR, Peng Z, Hill DP, and Aiken C (2011) Triggered creep as a possible mechanism for delayed dynamic triggering of tremor and earthquakes. *Nature Geoscience* 4: 384–388. http://dx.doi.org/10.1038/NGE0114.

Shimomura Y, Nishimura T, and Sato H (2006) Bubble growth processes in magma surrounded by an elastic medium. *Journal of Volcanology and Geothermal Research* 155: 307–322.

Sibson RH (1982) Fault zone models, heat flow, and the depth distribution of earthquakes in the continental crust of the United States. *Bulletin of the Seismological Society of America* 72: 151–163.

Singh SK, Anderson JG, and Rodriguez M (1998) Triggered seismicity in the Valley of Mexico from major Mexican earthquakes. *Geofisica Internacional* 37: 3–15.

Spudich P, Steck LK, Hellweg M, Fletcher JB, and Baker LM (1995) Transient stresses at Parkfield, California, produced by the M 7.4 Landers earthquake of June 28, 1992: Observations from the UPSAR dense seismograph array. *Journal of Geophysical Research* 100: 675–690.

Stark MA and Davis SD (1995) Remotely triggered microearthquakes at The Geysers geothermal field. *California Geophysical Research Letters* 23: 945–948.

Steacy S, Gomberg J, and Cocco M (2005) Introduction to special section: Stress transfer, earthquake triggering, and time-dependent seismic hazard. *Journal of Geophysical Research* 110. http://dx.doi.org/10.1029/2005JB003692.

Steeples DW and Steeples DD (1996) Far field aftershocks of the 1906 earthquake. *Bulletin of the Seismological Society of America* 86: 921–924.

Stein RS (1999) The role of stress transfer in earthquake occurrence. *Nature* 402: 605–609.

Stein RS and Lisowski M (1983) The 1979 Homestead Valley earthquake sequence, California: Control of aftershocks and postseismic deformation. *Journal of Geophysical Research* 88: 6477–6490.

Sturtevant B, Kanamori H, and Brodsky E (1996) Seismic triggering by rectified diffusion in geothermal systems. *Journal of Geophysical Research* 101: 25269–25282.

Surve G and Mohan G (2012) Possible evidence of remotely triggered and delayed seismicity due to the 2001 Bhuj earthquake ($M_w = 7.6$) in western India. *Natural Hazards* 64: 299–310.

Tanaka S, Ohtake M, and Sato H (2004) Tidal triggering of earthquakes in Japan related to the regional tectonic stress. *Earth, Planets and Space* 56: 511–515.

Tang C-C, Peng Z, Chao K, Chen C-H, and Lin C-H (2010) Detecting low-frequency earthquakes within the non-volcanic tremor in southern Taiwan triggered by the 2005 M_w 8.6 Nias earthquake. *Geophysical Research Letters* 37: L16307.

Tang C-C, Peng Z, Lin C-H, Chao K, and Chen C-H (2013) Statistical properties of low-frequency earthquakes triggered by large earthquakes in Southern Taiwan. *Earth and Planetary Science Letters* 373: 1–7, j.epsl.2013.04.039.

Tape C, West M, Silwal V, and Ruppert N (2013) Earthquake nucleation and triggering on an optimally oriented fault. *Earth and Planetary Science Letters* 363: 231–241.

Thomas AM, Burgmann R, Shelly DR, Beeler NM, and Rudolph ML (2012) Tidal triggering of low frequency earthquakes near Parkfield, California: Implications for fault mechanics within the brittle-ductile transition. *Journal of Geophysical Research* 117: B05301. http://dx.doi.org/10.1029/JB009036.

Thomas AM, Nadeau RM, and Burgmann R (2009) Tremor-tide correlations and near-lithostatic pore pressure on the deep San Andreas fault. *Nature* 462: 1048–1052.

Tibi R, Wiens DA, and Inoue H (2003) Remote triggering of deep earthquakes in the 2002 Tonga sequence. *Nature* 424: 921–925.

Toda S, Stein R, Beroza G, and Marsan D (2012) Aftershocks halted by static stress shadows. *Nature Geoscience* 5(6): 410–413.

Toda S, Stein RS, and Lin J (2011) Widespread seismicity excitation throughout central Japan following the 2011 M = 9.0 Tohoku earthquake and its interpretation by

Coulomb stress transfer. *Geophysical Research Letters* 38: L00G03. http://dx.doi.org/10.1029/2011GL047834.

Tolstoy M, Vernon FL, Orcutt JA, and Wyatt FK (2002) Breathing of the seafloor: Tidal correlations of seismicity at Axial volcano. *Geology* 30: 503–506.

Ukawa M, Fujita E, and Kumagai T (2002) Remote triggering of microearthquakes at the Iwo-Jima volcano. *Journal of Geography* 111: 277–286.

van der Elst NJ and Brodsky EE (2010) Connecting near-field and far-field earthquake triggering to dynamic strain. *Journal of Geophysical Research* 115. http://dx.doi.org/10.1029/2009JB006681.

van der Elst NJ, Brodsky EE, Le Bas P-Y, and Johnson PA (2012) Acoustic compaction in steady shear flow: Experimental evidence for suppression of shear dilatancy by internal acoustic vibrations. *Journal of Geophysical Research* 117: B09314. http://dx.doi.org/10.1029/2011B008927.

Velasco AA, Hernandez S, Parsons T, and Pankow K (2008) Global ubiquity of dynamic earthquake triggering. *Nature Geoscience* 1: 375–379.

Vohis RC (1967) *Hydrologic Effects of the Earthquake of March 27, 1964, Outside Alaska*. Washington, DC: US Geological Survey.

Voisin C, Campillo M, Ionescu IR, Cotton F, and Scotti O (2000) Dynamic versus static stress triggering and friction parameters: Inferences from the November 23, 1980, Irpinia earthquake. *Journal of Geophysical Research* 105: 21647–21658.

Voisin C, Cotton F, and Di Carli S (2004) A unified model for dynamic and static stress triggering of aftershocks, antishocks, remote seismicity, creep events, and multisegmented rupture. *Journal of Geophysical Research* 109. http://dx.doi.org/10.1029/2003JB002886.

Walter TR (2007) How a tectonic earthquake may wake up volcanoes: Stress transfer during the 1996 earthquake-eruption sequence at the Karymsky volcanic group, Kamchatka. *Earth and Planetary Science Letters* 264: 347–359.

Walter TR, Wang R, Acocella V, Neri M, Grosser H, and Zschau J (2009) Simultaneous magma and gas eruptions at three volcanoes in southern Italy: An earthquake trigger? *Geology* 37: 251–254.

Wang T-H, Cochran ES, Agnew D, and Oglesby DD (2013) Infrequent triggering of tremor along the San Jacinto Fault near Anza, California. *Bulletin of the Seismological Society of America* 102: 2482–2497.

Wang C-Y and Manga M (2009) Hydrologic response to earthquakes and a general metric. *Geofluids* 10: 206–216.

Wang C-Y and Manga M (2010) *Earthquakes and Water*, 218 pp. Springer.

Wech AG, Boose CM, Stern TA, and Tounsend J (2012) Tectonic tremor and deep slow slip on the Alpine Fault. *Geophysical Research Letters* 39: L10303. http://dx.doi.org/10.1029/2012GL051751.

Wen K-L, Beresnev IA, and Cheng S (1996) Moderate-magnitude seismicity remotely triggered in the Taiwan region by large earthquakes around the Philippine Sea Plate. *Bulletin of the Seismological Society of America* 86: 843–847.

West M, Sanchez JJ, and McNutt SR (2005) Periodically triggered seismicity at Mount Wrangell, Alaska, after the Sumatra earthquake. *Science* 308: 1144–1146.

Wiemer S (2001) A software package to analyze seismicity: ZMAP. *Seismological Research Letters* 72: 373–380.

Wilcock WS (2001) Tidal triggering of microearthquakes on the Juan de Fuca Ridge. *Geophysical Research Letters* 28: 3999–4002.

Wu P and Johnston P (2000) Can deglaciation trigger earthquakes in North America? *Geophysical Research Letters* 27: 1323–1326.

Wu C, Peng Z, Wang W, and Chen Q-F (2011) Dynamic triggering of shallow earthquakes near Beijing, China. *Geophysical Journal International* 185: 1321–1334.

Wu J, Peng Z, Wang W, Gong X, Chen Q, and Wu C (2012) Comparisons of dynamic triggering near Beijing, China following the recent Sumatra earthquakes. *Geophysical Research Letters* 39: L21310. http://dx.doi.org/10.1029/2012GL053515.

Yamazaki K, Teraishi M, Komatsu S, Sonoda Y, and Kano Y (2011) On the possibility of the 2011 Tohoku-oki earthquake reactivating Shinmoe-dake volcano, southwest Japan: Insight from strain data measured in vaults. *Natural Hazards and Earth System Science* 11: 2655–2661.

Yang H and Peng Z (2013) Lack of additional triggered tectonic tremor around the Simi Valley and the San Gabriel Mountain in Southern California. *Bulletin of the Seismological Society of America* 103: 3372–3378.

Yukutake Y, Honda R, Harada M, Aketagawa T, Ito H, and Yoshida A (2011) Remotely triggered seismicity in the Hakone volcano following the 2011 off the Pacific coast of Tohoku earthquake. *Earth, Planets and Space* 63: 737–740. http://dx.doi.org/10.5047/eps.2011.05.004.

Yukutake Y, Miyazawa M, Honda R, et al. (2013) Remotely triggered seismic activity in Hakone volcano during and after the passage of surface waves from the 2011 M 9.0 Tohoku-Oki earthquake. *Earth and Planetary Science Letters* 373(1): 205–216.

Zigone D, Rivet D, Radiguest M, et al. (2012) Triggering of tremors and slow slip event in Guerrero, Mexico, by the 2010 Mw 8.8 Maule, Chile, earthquake. *Journal of Geophysical Research* 117: B09304. http://dx.doi.org/10.1029/1012JB009160.

Ziv A (2006) On the role of multiple interactions in remote aftershock triggering: The Landers and the Hector Mine case studies. *Bulletin of the Seismological Society of America* 96: 80–89.

Zoback MD and Zoback ML (2002) State of stress in the Earth's lithosphere. In: Lee WHK, Kanamori H, Jennings PC, and Kisslinger C (eds.) *International Handbook of Earthquake and Engineering Seismology, Part A*, pp. 559–568. Amsterdam: Academic Press.

4.12 Earthquake Hydrology

M Manga and C-Y Wang, University of California, Berkeley, CA, USA

© 2015 Elsevier B.V. All rights reserved.

4.12.1	**Introduction**	305
4.12.2	**Hydrologic Response to Stress**	305
4.12.2.1	Effect of Static Stress	306
4.12.2.1.1	Poroelastic flow and deformation	306
4.12.2.1.2	Permanent deformation	307
4.12.2.2	Effect of Dynamic Strain	308
4.12.2.2.1	Poroelastic deformation and fluid flow	308
4.12.2.2.2	Permanent deformation	308
4.12.2.2.3	Liquefaction	308
4.12.3	**Observations and Their Explanations**	310
4.12.3.1	Wells	310
4.12.3.1.1	Water-level oscillations	310
4.12.3.1.2	Persistent and delayed postseismic changes in water level in the near and intermediate field	310
4.12.3.1.3	Near-field response of unconsolidated materials	312
4.12.3.1.4	Intermediate-field and far-field response	314
4.12.3.1.5	Summary	315
4.12.3.2	Streamflow	315
4.12.3.3	Mud Volcanoes	319
4.12.3.4	Geysers	320
4.12.4	**Feedback Between Earthquakes and Hydrology**	321
4.12.5	**Hydrologic Precursors**	322
4.12.6	**Concluding Remarks**	323
Acknowledgments		324
References		324

4.12.1 Introduction

For thousands of years, changes in the amount, direction, and rate of water flow in streams and in fluid pressure in the subsurface have been documented following large earthquakes. Examples include the formation of new springs, disappearance of previously active springs, increased discharge in streams by up to an order of magnitude that persists for weeks or even months, fluctuations in water levels in wells by several meters, and permanent changes in water levels in wells coincident with earthquakes located thousands of kilometers away. Such observations, although often simply assigned to the category of curious phenomena, provide unique insight into hydrogeologic processes and, in particular, their correlation with tectonic processes at spatial and temporal scales that could otherwise not be studied.

Hydrologic responses are not unexpected: Earthquakes cause strain, and strain changes fluid pressure and alters hydrogeologic properties such as permeability, which controls the rate of fluid flow. What is more surprising about many hydrologic responses, however, is their large amplitude and the great distances over which they occur.

This chapter is divided into two main parts. First, we describe the processes by which fluid pressure and hydrologic properties can change following an earthquake. Next, we review hydrologic phenomena attributed to earthquakes and summarize our current state of understanding about the processes that cause these phenomena. We then conclude with a brief discussion of possible hydrologic precursors to earthquakes.

4.12.2 Hydrologic Response to Stress

Earthquakes change the static stress (i.e., the offset of the fault generates a static change in stress in the crust) and also cause dynamic stresses (from the seismic waves). Both stresses increase as the seismic moment of the earthquake, but they decay very differently with distance r. Static stresses decrease as $1/r^3$. By comparison, dynamic stresses, which are proportional to the seismic wave amplitude, decrease more gradually. A standard empirical relationship between surface wave amplitude and magnitude has dynamic stress decreasing as $1/r^{1.66}$ (e.g., Lay and Wallace, 1995). Other factors such as directivity, radiation pattern, and crustal structure will also influence the amplitude of shaking. However, the significant difference in the dependence on distance is a robust feature distinguishing static and dynamic stresses.

Table 1, reproduced from Manga and Brodsky (2006), lists and compares the magnitude of earthquake stresses with other external stresses of hydrologic interest. Both static and dynamic stresses may be significant in the near field and intermediate field (within up to a few fault lengths), but only dynamic

 http://dx.doi.org/10.1016/B978-0-444-53802-4.00082-8

Table 1 Amplitude and period of stress changes of earthquakes compared with other sources of stress

	Stress (MPa)			*Period*
Solid earth tides	10^{-3}			12 h
Ocean tides	10^{-2}			12 h
Hydrologic loading	$10^{-3}-10^{-1}$			Days–years
Glacier loading	$10^{1}-10^{2}$			10^{3} years
Epicentral distance	10^{2} km	10^{3} km	10^{4} km	
Static stress changes, M8	10^{-1}	10^{-4}	10^{-7}	NA
Dynamic stress changes, M8	3	0.06	0.001	20 s

Adapted from Manga M and Brodsky EE (2006) Seismic triggering of eruptions in the far field: Volcanoes and geysers. *Annual Review of Earth and Planetary Sciences* 34: 263–291.

stresses are larger than tidal stress in the far field (many fault lengths away).

4.12.2.1 Effect of Static Stress

4.12.2.1.1 Poroelastic flow and deformation

In response to stresses, porous solids deform, the pressure of fluids within pores changes, and pore fluids can flow. The basic theory of poroelasticity, first developed by Biot (1941, 1956a,b), describes the coupling between changes in stress σ, strain ϵ, pore pressure p, and the material properties that relate these three variables. A complete discussion and derivation of the theory, even its simplest forms, is beyond the scope of this chapter, but several key relationships are useful for understanding the hydrologic response to earthquakes and interpreting observations. Wang's (2000) text on linear poroelasticity and review papers by Rice and Cleary (1976), Roeloffs (1996), and Neuzil (2003) provide thorough and pedagogical reviews.

Consider an isothermal, linear, isotropic poroelastic material subjected to a stress σ_{ij}. The material will respond by deforming with the total strain ϵ_{ij}, given by

$$\epsilon_{ij}=\frac{1}{2G}\left[\sigma_{ij}-\frac{v}{1+v}\sigma_{kk}\delta_{ij}\right]+\frac{\alpha}{3K}p\delta_{ij} \qquad [1]$$

where G is the shear modulus, K is the bulk modulus, v is the Poisson ratio, and α is the Biot–Willis coefficient, which is the ratio of the increment in fluid content to the volumetric strain at constant pore pressure. The increment in fluid content, denoted ψ, is the change in fluid mass per unit volume divided by the density of the fluid at the reference state and is given by

$$\psi=\frac{\alpha}{3K}\sigma_{kk}+\frac{\alpha}{KB}p \qquad [2]$$

where B is called Skempton's coefficient and has a value between 0 and 1, with 'hard rocks' such as sandstone and granite having values between about 0.5 and 0.9 and unconsolidated materials having a value close to 1. The variable B is related to the porosity ϕ and to the compressibilities of the pore fluid, solid grains, and saturated rock. The sign convention is chosen so that negative σ_{kk} indicates compression.

First, we consider the so-called 'undrained' limit in which there is no flow of pore fluid, that is, $\psi=0$ (the opposite extreme, termed 'drained,' corresponds to the limit in which there is no change in fluid pressure, i.e., $p=0$). The change in pore pressure p caused by a change in mean stress σ_{kk} can be found from eqn [2]:

$$p=-\frac{B}{3}\sigma_{kk} \qquad [3]$$

Equation [3] can be expressed in terms of the volumetric strain ϵ_{kk} using constitutive relationship [1] and physical relationship [2]:

$$p=-\frac{B}{3}\sigma_{kk}=-BK_u\epsilon_{kk}=-\frac{2GB}{3}\frac{1+v_u}{1-2v_u}\epsilon_{kk} \qquad [4]$$

where the subscript u indicates a value under undrained conditions. Shear modulus G should have the same value in drained and undrained conditions for a linear poroelastic material.

If there are gradients of fluid pressure, there will be a flux of fluid $\mathbf{q}$ (volume/area per unit time) governed by Darcy's equation

$$\mathbf{q}=-\frac{k}{\mu}\nabla p \qquad [5]$$

where the pressure p is the excess pressure, that is, the pressure difference from hydrostatic, k is the permeability, and μ is the fluid viscosity. p, as defined, is equivalent to the hydraulic head multiplied by ρg, where ρ is the density of the fluid and g the gravitational acceleration. Darcy's equation, coupled with the continuity equation governing conservation of mass, if the permeability k and μ are constant, leads to pore pressure being governed by an inhomogeneous diffusion equation

$$\frac{\alpha}{KB}\left[\frac{B}{3}\frac{\partial\sigma_{kk}}{\partial t}+\frac{\partial p}{\partial t}\right]=\frac{k}{\mu}\nabla^2 p \qquad [6]$$

The term involving the time derivative of stress couples the mechanical deformation and the fluid flow. The time derivative of stress will not appear in the case of uniaxial strain and constant vertical stress, both reasonable approximations for near-surface aquifers (but not rigorously correct if there is two- or three-dimensional flow), and can be neglected for very compressible fluids such as gases (Wang, 2000). In these cases, the evolution of fluid pressure resulting from flow is uncoupled from that for the stress and can be solved independently. Equation [6] then reduces to the form commonly used to model groundwater flow:

$$S\frac{\partial p}{\partial t}=\frac{k\rho g}{\mu}\nabla^2 p \quad \text{or} \quad \frac{\partial p}{\partial t}=D\nabla^2 p \qquad [7]$$

where S is standard hydrogeologic specific storage and is equal to volume of water released per unit change in head while maintaining no lateral strain and constant vertical stress and D is called the hydraulic diffusivity.

Assuming eqn [7] is a good approximation, the timescale τ for changes in pore pressure to diffuse or relax can be calculated by $\tau\sim L^2/D$, where L is the length scale over which the pressure diffuses. Typical values of D are 10^{-10}–$10^{4}\,\mathrm{m^2\,s^{-1}}$, with small values characterizing low-permeability rocks (e.g., shale) and large values characterizing coarse, unconsolidated sediments or highly permeable rock (e.g., gravel and karst, respectively). A value of $10^{0}\,\mathrm{m^2\,s^{-1}}$ would be typical of

sandstone aquifers. For a 100 m thick layer, then, pore pressure changes will diffuse on timescales that range from minutes to many decades and with a timescale of hours for a sandstone aquifer.

We will see, when we examine hydrologic observations, that many of the most dramatic hydrologic phenomena mentioned in the introduction, and in particular those that involve water discharge at the Earth's surface, cannot be explained by traditional linear poroelasticity theory. The reason is quite straightforward. Consider the case in which a stress change σ_{kk} causes an undrained pressure change p given by eqn [3]. For a (very large) static stress change of 1 MPa (see **Table 1**), eqn [3] implies pressure changes of 0.1–0.3 MPa. This would cause a 10–30 m change in water level in a well. To bring the pore pressure back to its original value requires the removal of an increment of fluid content

$$\psi = \frac{3(v_u - v)}{2GB(1+v)(1+v_u)}\sigma_{kk} \qquad [8]$$

Using values for Berea sandstone (Wang, 2000) of $v = 0.17$, $v_u = 0.33$, $G = 5.6$ GPa, and $B = 0.75$, we obtain $\psi = 3.5 \times 10^{-5}$. For a 100 m thick aquifer with porosity $\phi = 0.4$, this is equivalent to about 1 mm of water per unit area discharged at the surface – an amount that is small compared with that needed to explain the excess discharge in streams following large earthquakes (see Table 1 in Manga, 2001).

Equations [1]–[3] apply for linear, isotropic porous materials. In this limit, shear stresses cause no change in pore pressure. In natural rocks, however, shear stresses can induce changes in pore pressure because of anisotropy or the nonlinear elastic behavior of rocks (e.g., Hamiel et al., 2005).

4.12.2.1.2 Permanent deformation

At shear strains from 10^{-4} to 10^{-3}, brittle rocks, sediments, and soils show permanent deformation. At even greater shear strains ($\sim 10^{-2}$), failure occurs. The changes in porosity of sediments and soils during permanent deformation may be many orders of magnitude greater than that during elastic deformation. Thus, under undrained conditions, permanent deformation can be associated with significant changes in pore pressure and groundwater flow.

During permanent deformation, the cohesion among the grains of sediments and soils is disrupted and the grains tend to move to reach a new state of equilibrium. The basic characteristics of deformation, however, depend on the consolidation state of sediments and soils. For loose deposits, shear deformation causes grains to move into preexisting pores. This reduces the original porosity and thus the volume of sediments (soils) decreases – a process commonly known as consolidation. Shear deformation of dense deposits, in contrast, will cause sediment grains to roll over each other and, in so doing, create new porosity; thus, the volume of sediments (soils) increases – a process commonly known as dilatancy. In both cases, a critical state is eventually reached in undrained deformation, beyond which porosity no longer changes with continued shearing.

The deformation behavior of brittle rocks is distinctly different from that of sediments and soils, leading to different predictions for the hydrologic and seismic consequences under applied stresses. Laboratory measurements show that beyond the elastic limit, shearing of consolidated rocks causes microcracks to open and the volume of rock to increase – that is, they become dilatant (e.g., Brace et al., 1966). At still higher deviatoric stresses, microcracks may coalesce and localize into a shear zone, leading to eventual rupture (Lockner and Beeler, 2002). Repeated rupturing of the brittle crust along a fault zone can produce pulverized gouge material with significant porosity. Geologic investigations of exhumed fault zones in California (e.g., Chester et al., 1993, 2004) and in Japan (Forster et al., 2003; Wibberley and Shimamoto, 2003) and studies of drilled cores from fault zones (e.g., Ohtani et al., 2001; Zoback et al., 2010) have demonstrated that some fault zones are composed of fluid-saturated, porous fault gouge several hundred meters thick. Attempts to image the structure of fault zones by using conventional seismic reflection and refraction methods (e. g., Thurber et al., 1997) suffered from not having high enough station and source densities to resolve deep fault zone structures on the scale of hundreds of meters (Mizuno, 2003). However, seismic imaging of the fault zone structures with fault zone trapped waves (e.g., Li and Malin, 2008; Li et al., 2000; Mizuno and Kuwahara, 2008) and high-frequency body waves (Li and Zhu, 2007) has delineated a 100–250 m wide gouge zone along active faults in California (Li and Zhu, 2007; Li et al., 2000) and in Japan (e.g., Mizuno, 2003) in which the seismic velocities are reduced by 30–50% from the wall-rock velocities. Direct drilling of the San Andreas Fault zone to a depth of $\sim$3 km near Parkfield, CA, also revealed a process zone 250 m in width, in which the seismic velocities are 20–30% below the wall-rock velocities (Zoback et al., 2010).

Using high-resolution electromagnetic imaging, Unsworth et al. (2000) showed that the San Andreas Fault zone contains a wedge of low-resistivity material extending to several kilometers in depth; they further suggested a range of saturated porosity to account for the observed electrical resistivity. Using experimentally measured seismic velocity and electrical conductivity of fault gouge at elevated pressures to interpret the then-available seismic velocity and electrical resistivity of the San Andreas Fault zone in Central California, Wang (1984) suggested that the fault zone may consist of saturated fault gouge. Inverting the gravity anomaly across the San Andreas Fault zone in Central California and constraining the inversion by using existing seismic reflection profiles, Wang et al. (1986) further showed that the fault zone may be characterized by relatively low density with a corresponding porosity greater than 10% extending to seismogenic depths.

That fault zones may be porous and saturated at seismogenic depths has important implications for a host of issues related to the dynamics of faulting (e.g., Andrew, 2003; Brodsky and Kanamori, 2001; Lachenbruch, 1980; Mase and Smith, 1987; Sibson, 1973; Sleep and Blanpied, 1992; Wibberley and Shimamoto, 2005), the strength (or weakness) of faults (e.g., Lachenbruch and Sass, 1977; Mount and Suppe, 1987; Rice, 1992; Zoback et al., 1987), and the forces that move lithospheric plates (e.g., Bird, 1978). To obtain direct information on the pore pressure in fault zones, drilling into active faults has been attempted around the world in the past decades (San Andreas Fault Observatory at Depth, SAFOD, near Parkfield in Central California (Zoback et al., 2010), the

Nojima Fault Zone Probe in Japan, the Chelungpu Fault Drilling Project in Taiwan (see Brodsky et al., 2010, for a summary of the Nojima and Chelungpu projects), and an ongoing project to drill into the New Zealand South Island Alpine Fault, Gorman, 2011). The SAFOD drilling project is the best documented and has provided invaluable information about the internal structure of the San Andreas Fault zone (e.g., Zoback et al., 2010). At a depth of ~2.7 km, the SAFOD drilling encountered three actively slipping, clay-rich gouge zones, each 23 m wide. Three observations made during drilling led Zoback et al. (2010) to conclude that there was no evidence of high pore pressure in the fault cores: (1) no excess influx of aqueous formation fluids from the fault zone into the borehole, (2) no excess inflow of formation gas into the borehole, and (3) a near-constant V_p/V_s ratio across the fault zone (where V_p and V_s are the velocities of the compressional and the shear seismic waves, respectively). However, quantitative simulation of the expected flow of aqueous fluids and gases into the well during drilling and the V_p/V_s ratios normal to the drilling direction (Wang, 2011), based on laboratory data for the permeability of fault zone material (Chu et al., 1981; Marone et al., 2010; Morrow et al., 1984; Rathbun et al., 2010) and the seismic wave velocities in layered rocks (Lo et al., 1986), showed that the time duration during drilling may not be long enough to detect any flow of aqueous fluids and gases into the well and that the V_p/V_s ratio measured may not be sensitive enough to determine whether high pore pressure does exist in the fault zone. Thus, long-term monitoring of pore pressure may be required to provide definite knowledge of whether high pore pressure does or does not exist in fault zones.

4.12.2.2 Effect of Dynamic Strain

4.12.2.2.1 Poroelastic deformation and fluid flow

Seismic waves cause spatial variations in strain and hence spatial variations in pore pressure. The resulting pore pressure gradients cause fluid flow. Seismic wave speeds in saturated rocks will thus differ from those in unsaturated rocks and will be frequency-dependent. As the fluid moves, energy is also dissipated by the fluid flow, which in turn is governed by the permeability of the porous material, the geometry of the pores, and the viscosity of the fluid. This contributes to seismic attenuation. Wave speed and attenuation can be calculated theoretically as a function of frequency for porous materials with different pore structures by solving the poroelastic equations coupled with Darcy's equation for fluid flow at the macroscopic scale. As frequency increases, pore scale (or 'squirt') flows become more important (e.g., Mavko et al., 1998). These ideas were first quantified in a series of papers by Biot (e.g., Biot, 1941, 1956a,b, 1962) and extended by many authors since, for example, by accounting for squirt flows (e.g., Dvorkin et al., 1994). Additional important extensions to actual geologic settings include accounting for heterogeneity with dual-porosity models (e.g., Berryman and Wang, 2000) or fractal geometries (e.g., Masson and Pride, 2006) and partial or multiphase saturation (e.g., Pride et al., 2004).

Changes in pore pressure and stress also change fracture and pore geometry, which in turn alters wave speeds and anisotropy. The frequency dependence of attenuation and wave speeds also depends on hydrogeologic properties such permeability and the nature of permeability heterogeneity. The relationship between seismic properties, hydrogeologic properties, and fluid pressure involves so many parameters that inferring particular properties of the subsurface from seismic measurements is challenging. However, measurements of seismic attenuation, seismic anisotropy, and wave speeds can be used as a noninvasive probe to monitor changes in the subsurface (e.g., Liu et al., 2004).

4.12.2.2.2 Permanent deformation

The dynamic deformation of sediments and soils, in contrast to the static deformation, is affected by inertial forces that depend on loading rates and number of loading cycles. An often-used measure of the total amount of seismic energy imparted to an engineered structure is the Arias intensity, which is proportional to the square of acceleration (Arias, 1970; Jennings, 2003). Therefore, even if the magnitude of stresses is small, inertial forces may significantly affect deformation at high frequencies. For this reason, it is necessary in earthquake engineering to examine the dynamic deformation of sediments and soils for dynamic strains as small as 10^{-6}. Permanent volumetric deformation, however, does not occur until the magnitude of shear strain reaches some threshold. Dobry (1989) summarized undrained experimental results in which excess pore pressure developed during the application of ten cycles of sinusoidal loads. He showed that, for different sands with a wide range of dry densities subjected to a wide range of effective confining pressures, pore pressure did not increase until the cyclic shear strain amplitude was increased to 10^{-4}. The increase in pore pressure indicates a decrease in pore volume, and hence consolidation. It is interesting to note that the strain threshold is comparable in magnitude to that for permanent deformation to occur in static deformation (**Section 4.12.2.1**).

The number of loading cycles during dynamic loading is another important factor in determining the mechanical responses of sediments and soils. Results of laboratory experiments show that during cyclic loading at constant strain amplitude, pore pressure in sediments and soils increases with the number of loading cycles. At decreasing strain amplitudes, a greater number of loading cycles is required to build up excess pore pressure to a given magnitude (Ishihara, 1996). During large earthquakes, about 10–20 cycles of major ground shaking occur, though the amplitude of each cycle will vary.

4.12.2.2.3 Liquefaction

When pore pressures approach the overburden pressure, soils lose their rigidity and become fluidlike – a phenomenon widely known as liquefaction (Seed and Lee, 1966; Terzaghi et al., 1996), which is a major source of seismic hazard for engineered structures. Liquefied sediments and soils in the subsurface are associated with substantial pore pressure, as indicated by the height of the smear on the wall in **Figure 1**, left by ejected sands during the 1999 M7.5 Chi-Chi earthquake in Taiwan. Most liquefaction occurs near the ruptured fault, but a large number of examples of liquefaction have been documented at distances beyond the near field (Wang, 2007).

Field and laboratory studies show that the occurrence of liquefaction depends on many factors, such as earthquake magnitude, shaking duration, peak ground motion, depth to the groundwater table, basin structures, site effects, and

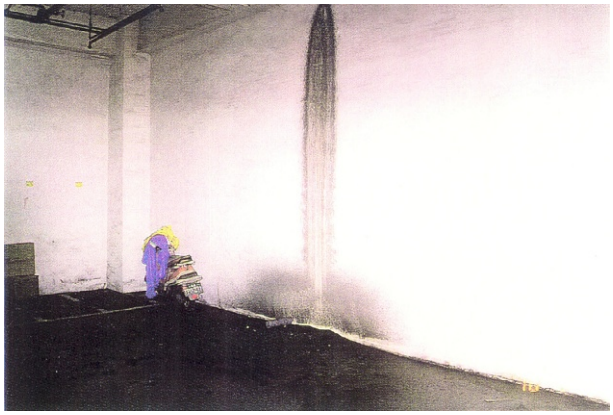


Figure 1 Smear of sediment left on a building wall created by ejected sand during the 1999 M7.5 Chi-Chi earthquake in Taiwan. Motorcycle at lower left shows scale. Reproduced from Su T-C, Chiang K-W, Lin S-J, Wang F-G, and Duann S-W (2000) Field reconnaissance and preliminary assessment of liquefaction in Yuan-Lin area. *Sino-Geotechnica* 77: 29–38.

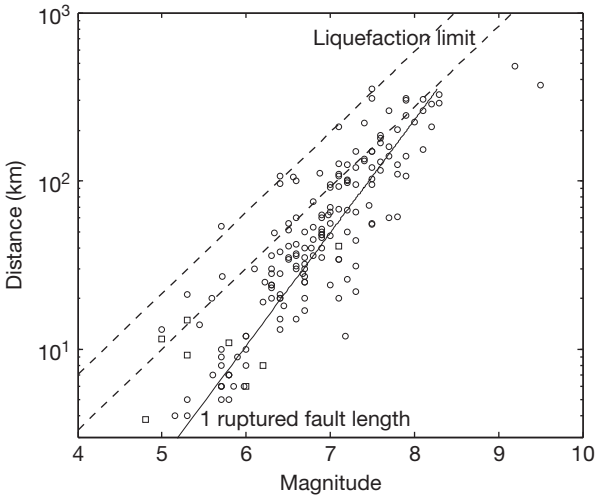


Figure 2 Relationship between earthquake magnitude (moment magnitude when known) and the distance between the earthquake epicenter and occurrence of liquefaction. Data shown with circles are from the compilation in Wang and Manga (2010a). Squares show data following the M7.1 2010 Darfield, New Zealand, earthquake and seven of its aftershocks (Simon Cox, GNS New Zealand, written communication) and responses to the M7.2 Fiordland, New Zealand, earthquake of 2003 (Hancox et al., 2003). For most of the New Zealand events, only the maximum distance for each earthquake is plotted. The dashed lines delineate an upper bound for the maximum distance over which shallow liquefaction has been observed and correspond to estimates of dissipated seismic energy density between 0.1 and 1 J m^{-3} (Wang, 2007). The solid line indicates one ruptured fault length based on the regression of Wells and Coppersmith (1994) and bounds the near field.

liquefaction susceptibility of sediments (e.g., Youd, 2003). Thus, the occurrence of liquefaction is difficult to predict on a theoretical basis, and empirical approaches are as a rule adopted in assessing the liquefaction potential of an area. Various ground penetration tests (e.g., Bardet, 2003) are routinely done during the geotechnical assessment of sites. On large scales, empirical relations between the magnitude of ground shaking and liquefaction (e.g., Wang et al., 2003) may be combined with numerical simulations of ground shaking during an earthquake (e.g., Stidham et al., 1999) to evaluate the regional liquefaction potential. Such applications, however, require extensive information about sediment properties and subsurface structures of the area of interest. In areas where such data are not available, a simpler approach may be applied to set some limits to the expected extent of liquefaction during potential earthquakes. Field observations show that, for earthquakes of a given magnitude M, the occurrence of liquefaction is confined within a particular distance from the earthquake epicenter, that is, the liquefaction limit R_{max} beyond which liquefaction is not expected (Ambraseys, 1988; Galli, 2000; Kuribayashi and Tatsuoka, 1975; Papadopoulos and Lefkopulos, 1993; Wang et al., 2005a; Wang, 2007). **Figure 2** shows a compilation of data for the occurrence of liquefaction and an empirical bound on the relationship between R_{max} and M (Wang, 2007); all data and references are tabulated in Wang and Manga (2010a) with additional data from the M7.1 2010 Darfield, New Zealand, earthquake and its aftershocks (Simon Cox, GNS, New Zealand, written communication) and responses to the M7.2 Fiordland, New Zealand, earthquake (Hancox et al., 2003). The empirical bounds correspond to estimates of dissipated seismic energy density e between 0.1 and 1 J m^{-3} (Wang and Manga, 2010b). e is defined as

$$e = \frac{1}{2}\sum_i \frac{\rho}{T_i}\int v_i(t)^2 \mathrm{d}t \qquad [9]$$

where the sum is taken over all modes of ground motion, ρ is density, and T_i and v_i are the period and velocity of the ith mode, respectively. Also shown for reference in **Figure 2** is the length of the ruptured fault as a function of M, based on the regression of Wells and Coppersmith (1994). While the data plotted in **Figure 2** neglect important features such as site effects and directivity, they do indicate that liquefaction occurs at distances beyond the near field (Wang, 2007). The liquefied sites near the liquefaction limit are likely to be those with the optimal conditions for liquefaction, that is, saturated soils with high liquefaction susceptibility or strong directivity effects. Liquefaction at closer distances may include sites with less optimal conditions that are exposed to greater seismic input.

Because liquefaction damages engineered structures, there has been a concerted research effort, both in the laboratory and in the field, to understand the processes of liquefaction and to predict its occurrences (for an overview of these efforts, see Bardet, 2003; National Research Council, 1985; Youd, 2003). In laboratory studies, sediment and soil samples are often obtained from the field and subjected to cyclic loading. At constant amplitude of deviatoric stress (stress minus the mean normal stress), deformation may be stable up to some number of cycles of shearing, but sudden onset of large shear deformation (onset of liquefaction) may occur if the cycles of shearing exceed some threshold number (e.g., National Research Council, 1985). The onset of liquefaction occurs once the effective stress (defined later in eqn [16]), which decreases with increasing pore pressure, permits the deviatoric stress to intersect the failure envelope of sediments and soils.

Laboratory studies have helped to advance our understanding of the processes of liquefaction and to identify the major physical factors that affect liquefaction. However, it is difficult to duplicate natural conditions in the laboratory or to obtain undisturbed samples from the field. Thus, laboratory studies may have limited practical application. In the field, tests such as the 'standard' penetration test are employed to assess the liquefaction potential of specific sites. Despite their wide application, these tests are empirical. Thus, much progress can still be made toward a better physical understanding of the liquefaction process.

4.12.3 Observations and Their Explanations

Here, we review hydrologic phenomena that accompany earthquakes. We will see that the magnitudes of hydrologic changes caused by earthquakes are typically small enough to be of little human consequence. This is not always the case; however, depletion of groundwater induced by seismic activity has been invoked to explain the abandonment of regions of Crete, Greece, during the Late Minoan period (Gorokhovich, 2005), and mud eruptions, possibly triggered by earthquakes, can be devastating (Mazzini et al., 2007).

We categorize observations by type of observation (wells, streamflow, mud volcanoes, and geysers). It will also be helpful when discussing responses to distinguish hydrologic responses by the spatial relationship between the observation and the earthquake. We will use the terms near field, intermediate field, and far field for distances within one fault length, up to several fault lengths, and many fault lengths, respectively. In the near field, changes in the properties of the fault zone itself may be responsible for hydrologic responses. Examples include the formation of new springs along ruptured faults (e.g., Lawson, 1908) or changes in groundwater flow paths because of changes in the permeability of, or near, the fault zone (e.g., Gudmundsson, 2000). In both the near and intermediate fields, dynamic and static strains are large enough that they might cause a measurable hydrologic response. In the far field, however, only dynamic strains are sufficiently large to cause responses. In this chapter, we focus on hydrologic responses outside the fault zone.

Hydrologic changes following earthquakes modify pore pressures, which in turn can influence local seismicity. The topic of triggered seismicity is covered in the chapter by Hill and Prejean (**Chapter 4.11**), and we simply note here that the mechanisms responsible for the observed hydrologic phenomena discussed next may be connected to those responsible for distant, triggered seismicity.

4.12.3.1 Wells

The water level in wells measures the fluid pressure at depths where the well is open to the surrounding formations. Several types of coseismic and postseismic responses are observed: water-level oscillations, coseismic changes in water level, and delayed changes in water level. **Figure 3** shows a particularly dramatic example of a response in a well in which water erupted to a height of 60 m above the surface following the 2004 M9.2 Sumatra earthquake 3200 km away.

Figure 3 Well in China responding to the 2004 M9.2 Sumatra earthquake 3200 km away. The picture was taken by Hou Banghua, Earthquake Office of Meizhou County, Guangdong, 2 days after the Sumatra earthquake. The fountain was 50–60 m high when it was first sighted 1 day after the earthquake.

4.12.3.1.1 Water-level oscillations

Water wells can act like seismometers by amplifying ground motions, in particular long-period Rayleigh waves. Water-level fluctuations as large as 6 m (peak-to-peak amplitude) were recorded in Florida, thousands of kilometers away from the epicenter, during the 1964 Alaska earthquake (Cooper et al., 1965). Hydroseismograms have been recorded since the early days of seismometer use (Blanchard and Byerly, 1935).

Figure 4 (from Brodsky et al., 2003) shows water level in a well in Grants Pass, Oregon, following the 2002 M7.9 Denali earthquake 3100 km away. Also shown is the vertical component of ground velocity measured on a broadband seismometer located adjacent to the well. Large water-level fluctuations, such as those shown in **Figure 4**, are unusual and occur only for the right combination of geometric (water depth in the well, well radius, and aquifer thickness) and hydrologic properties (transmissivity). The frequency-dependent response of the water level in a well to dynamic strain in an aquifer can be determined theoretically by solving the coupled equations for pressure change in the aquifer, flow toward/away from the well, and flow within the well (e.g., Cooper et al., 1965; Liu et al., 1989). In general, high transmissivity favors large amplitudes.

Kono and Yanagidani (2006) found that in closed (as opposed to open) borehole wells, pore pressure variations are consistent with an undrained poroelastic response (see **Section 4.12.2.1**) at seismic frequencies. In addition, they found no pore pressure response to shear deformation.

4.12.3.1.2 Persistent and delayed postseismic changes in water level in the near and intermediate field

Persistent changes in water level in wells are probably the best-documented hydrologic responses to earthquakes. Significant advances in analysis of water-level changes following earthquakes have been made during the last decade by compiling observations, identifying patterns in responses, and using quantitative models to interpret the observations

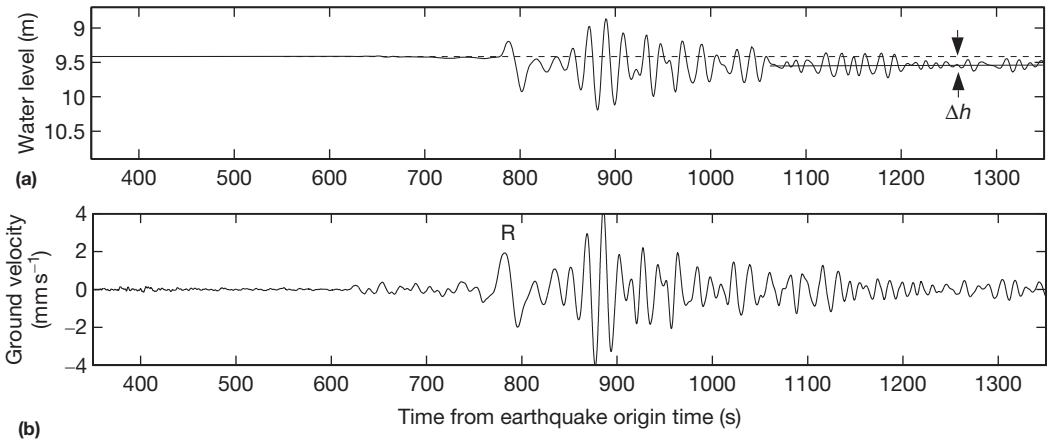


Figure 4 (a) Hydroseismograph recorded at 1 s intervals and (b) vertical component of ground velocity measured on a broadband seismometer at a well in Grants Pass, Oregon, following the 2002 M7.9 Denali earthquake located 3100 km away. Δh shows the 12 cm permanent change in water level that followed the passage of the seismic waves. *R* indicates the arrival of Rayleigh waves. Figure based on Figure 7 in Brodsky EE, Roeloffs E, Woodcock D, Gall I, and Manga M (2003) A mechanism for sustained ground water pressure changes induced by distant earthquakes. *Journal of Geophysical Research* 108: http://dx.doi.org/10.1029/2002JB002321.

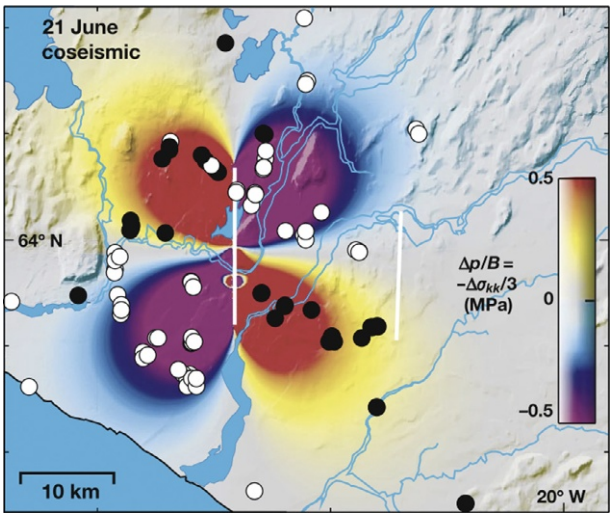


Figure 5 Predicted (color map) and observed (circles) coseismic water-level changes following a 2000 M6.5 strike-slip earthquake in Iceland. Black and white circles indicate water-level increases and decreases, respectively. The white line shows the mapped surface rupture. Figure from Jonsson S, Segall P, Pedersen R, and Bjornsson G (2003) Post-earthquake ground movements correlated to pore-pressure transients. *Nature* 424: 179–183.

(e.g., Brodsky et al., 2003; Chia et al., 2008; Cox et al., 2012; King et al., 1999; Kitagawa et al., 2006; Matsumoto and Roeloffs, 2003; Montgomery and Manga, 2003; Roeloffs, 1998; Roeloffs et al., 2003; Sato et al., 2004; Sil and Freymueller, 2006; Wang et al., 2003).

Sustained changes in water level in the near and possibly intermediate field can in some cases be explained by the coseismic static strain created by the earthquake. In this case, the water level will rise in zones of contraction and fall in regions of dilation.

Figure 5 (from Jonsson et al., 2003) shows a pattern of water-level change that mimics the pattern of coseismic volumetric strain after a strike-slip event in Iceland. An analogous correlation of volumetric strain and water-level changes was found after the 2003 M8.0 Tokachi-Oki thrust event in Japan (Akita and Matsumoto, 2004). Similar conclusions are reported by others (e.g., Quilty and Roeloffs, 1997; Wakita, 1975). This pattern of water-level change is not, however, universal. Following the 1999 M7.5 Chi-Chi earthquake in Taiwan, the water level rose in most near-field wells where the coseismic strain should have caused aquifer dilation (Koizumi et al., 2004; Wang et al., 2001). Following the 2004 M9 Sumatra earthquake, more than 5000 km away in Japan, Kitagawa et al. (2006) found that only about half the monitoring wells equipped with strain instruments recorded water-level changes consistent with the measured coseismic strain, implying that dynamic strains can also change water levels in the intermediate field.

Both the pattern and magnitude provide insight into the origin of water-level changes. The magnitude of expected water-level changes caused by coseismic strain can be determined from models of coseismic strain and models for well sensitivity that are calibrated on the basis of their response to barometric and tidal strains (e.g., Roeloffs, 1996). In some instances, recorded changes are similar to those predicted (e.g., Igarashi and Wakita, 1991), but changes can also be much larger than predicted, even in the near field (e.g., Igarashi and Wakita, 1995).

The time evolution of water-level changes provides additional constraints on the location and magnitude of pore pressure changes. **Figure 6** (from Roeloffs, 1998) shows water level for a period of 10 days following the 1992 M7.3 Landers earthquake 433 km from the well. The observed water-level changes, shown with data points, can be modeled by a coseismic, localized pore pressure change at some distance from the well (as shown by the solid curve). The origin, however, of the hypothesized pressure changes cannot be determined from measurements at the well (Roeloffs, 1998). Other processes that can cause delayed or gradual changes in water level are localized changes in porosity and/or permeability (e.g., Gavrilenko et al., 2000).

4.12.3.1.3 Near-field response of unconsolidated materials

At shallower depths and in unconsolidated materials, changes in water level in wells often exhibit larger amplitudes and different signs than predicted by the coseismic elastic strain. The 1999 Chi-Chi earthquake in Taiwan and, more recently, the 2010 Darfield earthquake in New Zealand provide excellent opportunities to examine models for postseismic water-level changes because the density of monitored wells was so high. We discuss these observations in much more detail than other types of well responses because the near-field hydrologic response at shallow depths and in unconsolidated materials is much less well studied. We focus, in our example, on the response to the Chi-Chi earthquake, because data for the New Zealand earthquakes are just appearing in the literature (Cox et al., 2012; Quigley et al., 2013) and there will undoubtedly be significant analyses appearing over the next few years.

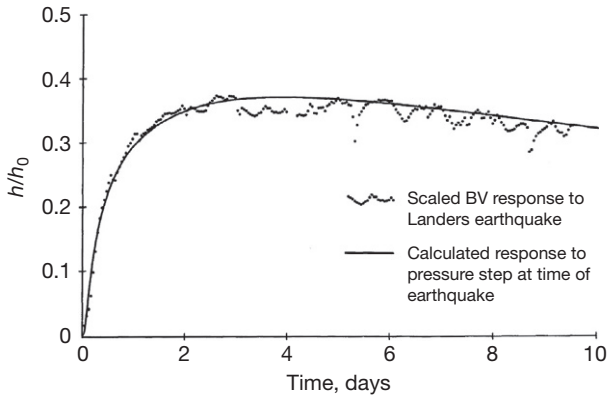


Figure 6 Water level as a function of time after the 1992 M7.2 Landers, CA, earthquake. The epicenter is 433 km away from the well. The smooth curve shows the predicted evolution of water level for a coseismic, localized pressure change at some distance from the well. h_0 is the magnitude of the head change where pore pressure changed. Figure from Roeloffs EA (1998) Persistent water level changes in a well near Parkfield, California, due to local and distant earthquakes. *Journal of Geophysical Research* 103: 869–889.

As a fortuitous coincidence, a network of 70 evenly distributed hydrologic stations (**Figure 7**), with a total of 188 wells, was installed in 1992 on a large alluvial fan (the Choshui River fan) near the Chi-Chi epicenter. The network was installed for the purpose of monitoring the groundwater resources over the Choshui River fan. At each station, one to five wells were drilled to depths ranging from 24 to 306 m to monitor groundwater level in individual aquifers. Each well was instrumented with a digital piezometer that automatically records the groundwater level hourly, with a precision of 1 mm (Hsu et al., 1999). The dense network of hydrologic stations and its close proximity to a large earthquake made the groundwater records during the Chi-Chi earthquake the most comprehensive and systematic so far obtained. In addition, a dense network of broadband strong-motion seismographic stations captured the ground motion in the area, and detailed studies of liquefaction occurrences were made. Finally, the isotopic composition of the groundwater was measured before and after the earthquake (Wang et al., 2005c).

The coseismic changes of groundwater level during this earthquake have been reported in several papers (Chia et al., 2001, 2008; Hsu et al., 1999; Lee et al., 2002; Wang et al., 2001). Near the ruptured fault, where consolidated sedimentary rocks occur near the surface, the groundwater level showed a coseismic, stepwise decrease (**Figure 8(a)** and **8(b)**). Away from the ruptured fault, in the fan of unconsolidated sediments, confined aquifers showed a coseismic stepwise rise (**Figure 8(c)** and **8(d)**). The magnitude of this rise increases with distance from the ruptured fault, reaching more than 5 m at distances of 20–30 km away from the fault before decreasing at greater distances (**Figure 7(a)**; Wang et al., 2001). In the uppermost aquifer, the coseismic changes in the groundwater

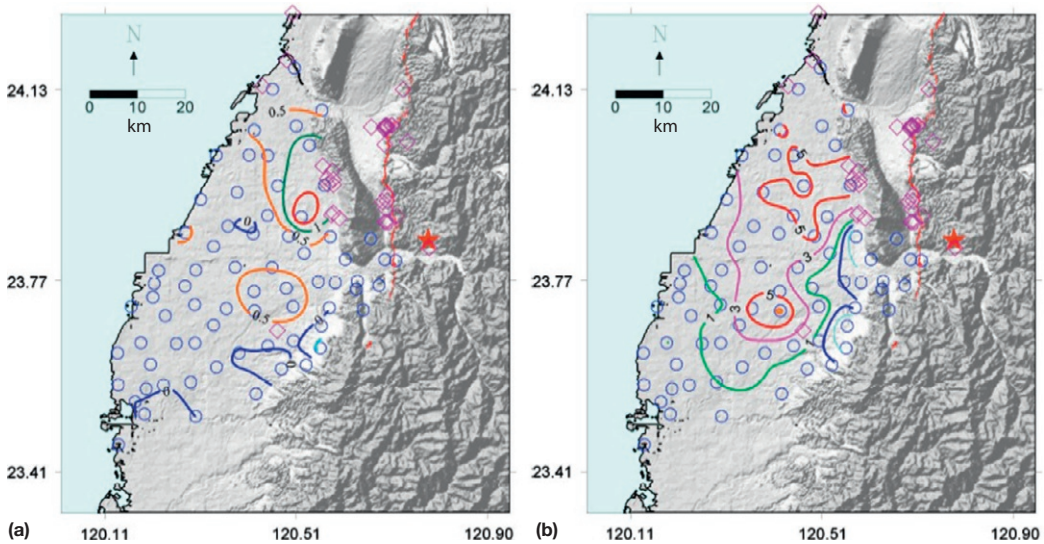


Figure 7 Spatial distribution of coseismic changes in groundwater level during the Chi-Chi earthquake. (a) Contours (m) showing coseismic groundwater-level changes in a confined aquifer (Aquifer II) in the Choshui River fan. (b) Contours (m) showing coseismic groundwater-level changes in the uppermost aquifer in the Choshui River fan. Groundwater monitoring stations are shown with open circles, liquefaction sites with open diamonds, epicenter of Chi-Chi earthquake with a star, and the ruptured fault with discontinuous red traces. Reproduced from Wang C-Y, Wang C-H, and Kuo C-Y (2004a) Temporal change in groundwater level following the 1999 (Mw = 7.5) Chi-Chi earthquake, Taiwan. *Geofluids* 4: 210–220.

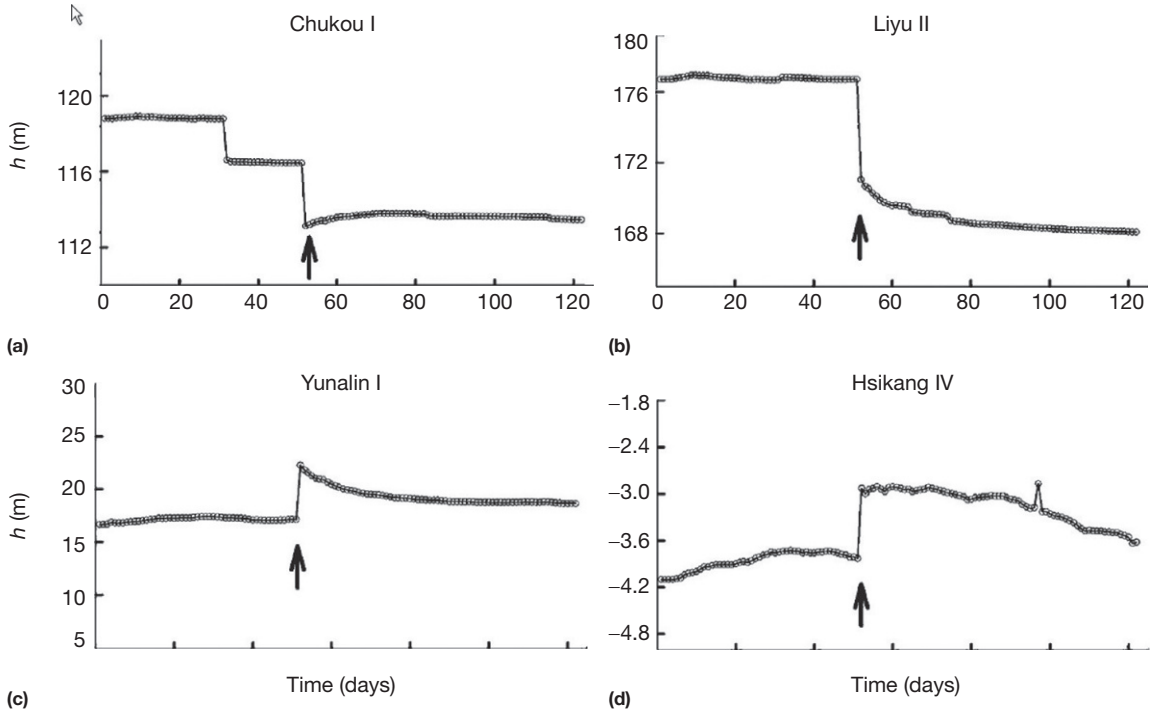


Figure 8 Four types of groundwater-level responses during the Chi-Chi earthquake. (a) Groundwater level dropped during the earthquake and gradually rose afterward. (b) Groundwater level dropped during the earthquake and continued to drop gradually afterward. (c) Groundwater level rose during the earthquake and gradually declined afterward. (d) Groundwater level rose during the earthquake and remained at the new level for some time before slowly declining. The different groundwater-level responses may be due to differences in the proximity to the ruptured fault and different hydrogeologic conditions as discussed in Wang et al. (2001). Figure created by Chung-Ho Wang.

level were much smaller in magnitude, with an irregular pattern (**Figure 7(b)**). Liquefaction sites on the Choshui River fan are clearly correlated with significant coseismic rise of the groundwater level in the uppermost aquifer; but no such correlation exists for the confined aquifers (**Figure 7**). This observation strongly indicates that liquefaction occurred in the uppermost aquifer but not in the confined aquifers. In the basins nearest the ruptured fault (east of the Choshui River fan), a comparison between liquefaction sites and the coseismic change in the groundwater level cannot be made, because no wells were installed and thus no groundwater-level data were available.

Lee et al. (2002) claimed that the spatial distribution of the groundwater-level change during the Chi-Chi earthquake may be accounted for by a poroelastic model. The widespread occurrence of liquefaction, however, provides clear evidence that at least part of the sediment deformation was not elastic. Furthermore, for confined aquifers that have a typical thickness of 1–10 m, the required changes in porosity would have to be from 10^{-3} to 10^{-2} to account for the range of coseismic groundwater-level changes (Wang et al., 2001), beyond the magnitude of strains calculated from an elastic model (Lee et al., 2002). Finally, the sign of the changes in water level is in general the opposite of those expected from a poroelastic model (Koizumi et al., 2004; Wang et al., 2001).

As noted earlier (**Section 4.12.2.2**), when loose sediments and soils are sheared at stresses between some upper and lower thresholds, grains tend to compact, leading to a decrease in porosity. If the deformation is undrained, as expected during seismic shaking, pore pressure would increase when the strain exceeds about 10^{-4}. This may explain the stepwise rise in the groundwater level in the fan (**Figure 8(c)** and **8(d)**). On the other hand, when consolidated sediments or sedimentary rocks are cyclically sheared beyond some critical threshold, microfracture and fracture may occur, leading to increased porosity and permeability and decreased pore pressure. This may explain the coseismic decline in the groundwater level in sedimentary rocks close to the ruptured fault (**Figure 8(a)** and **8(b)**).

It is possible to make an order-of-magnitude estimate of the change in porosity ϕ of the confined aquifer required to account for the coseismic water-level change. From the definition of specific storage $S = \rho^{-1}\partial(\rho\phi)/\partial h$ where h is the hydraulic head (Wang, 2000, p. 12), we have

$$\frac{\partial h}{\partial t} = \frac{1}{\rho S}\frac{\partial(\rho\phi)}{\partial t} \sim \frac{1}{S}\frac{\partial\phi}{\partial t} \qquad [10]$$

The term on the right-hand side of eqn [10] is the rate of change of sediment porosity at a given location resulting from ground shaking, which we assume is proportional to the intensity of ground shaking. Given that the intensity of ground shaking during the Chi-Chi earthquake decayed approximately exponentially with time (Ma et al., 1999), the rate of change in sediment porosity at a given location under earthquake shaking may be expressed as

$$\frac{d\phi}{dt} = Ae^{-\beta t} \qquad [11]$$

where β is on the order of $10^{-2} - 10^{-1}\,s^{-1}$ and A is a constant to be determined. Substituting eqn [11] into eqn [10] and

integrating with respect to time, we obtain the coseismic change in the hydraulic head:

$$\Delta h(t) = \frac{A}{S\beta}\left(e^{-\beta t} - 1\right) \qquad [12]$$

Pumping tests in the Choshui River fan yield $Sb \sim 10^{-3}$ (Lee and Wu, 1996) where b is the aquifer thickness. Given that ground shaking on the Choshui River fan during the Chi-Chi earthquake lasted between 20 and 100 s (Lee et al., 2002), for a 10 m groundwater-level change in a 1 m thick confined aquifer, $A = -10^{-4}$ to $-10^{-3}\,s^{-1}$, and for the same level of change in a 10 m aquifer, $A = -10^{-5}$ to $-10^{-4}\,s^{-1}$. We obtain a coseismic change in porosity between 10^{-3} and 10^{-2}.

The uppermost aquifer on the Choshui River fan is either unconfined or poorly confined, resulting in less regular changes in pore pressure and groundwater level. Thus, qualitatively, the processes of consolidation of sediments and dilation of sedimentary rocks, respectively, appear to provide a self-consistent interpretation of the spatial pattern of the coseismic groundwater-level changes on the Choshui River fan during the Chi-Chi earthquake (Wang et al., 2001).

The availability of both well records and strong-motion records at the time of the Chi-Chi earthquake allows quantitative examination of the relationship between ground motion, groundwater-level change, and the distribution of liquefaction sites. Interpolating the strong-motion parameters at the well sites, Wang et al. (2003) showed that there is only a weak correlation between the horizontal component of the peak ground acceleration (PGA) and the occurrence of water-level changes and liquefaction (**Figure 9(a)**). Wong and Wang (2007) confirmed this finding using a more extensive data set. This result is unexpected, because PGA is frequently used in engineering practice to predict the occurrence of liquefaction. Instead, Wong and Wang (2007) showed that there is a better correlation between the horizontal peak ground velocity (PGV) and the sites of liquefaction. Because PGV depends more on low-frequency components of ground motion than does PGA, this indicates that low-frequency ground motions are better correlated with elevated pore pressure and groundwater level than are high-frequency motions. This conclusion supports the finding that spectral accelerations and velocities at frequencies of about 1 Hz and lower were strongly correlated with the distribution of liquefaction sites (**Figure 9(b)**), whereas those above about 1 Hz were not (Wang et al., 2003; Wong and Wang, 2007). A similar conclusion was reached by comparing the hydrologic changes with various measures of earthquake intensity computed from seismograms passed through low-pass and high-pass filters (Wong and Wang, 2007). Although there is extensive evidence that low-frequency ground motions are better correlated with elevated pore pressure and groundwater level than high-frequency motions, it is unclear at this time whether the low-frequency ground motion was the cause for the coseismic groundwater-level change and liquefaction or whether the hydrologic changes, including liquefaction, caused the different spectral composition of the observed seismograms.

4.12.3.1.4 Intermediate-field and far-field response

In the far field (many fault lengths), the static stress due to the earthquake is nearly zero. Thus, sustained changes in water level at such distances (e.g., **Figure 4**) must be caused by the interactions between the aquifer and seismic waves. Such changes are not common, but occur at a small number of unusually sensitive sites usually located near active faults (King et al., 2006; Shi et al., 2014). The mechanism for these interactions is a matter of debate (Brodsky et al., 2003; Matsumoto and Roeloffs, 2003; Roeloffs, 1998).

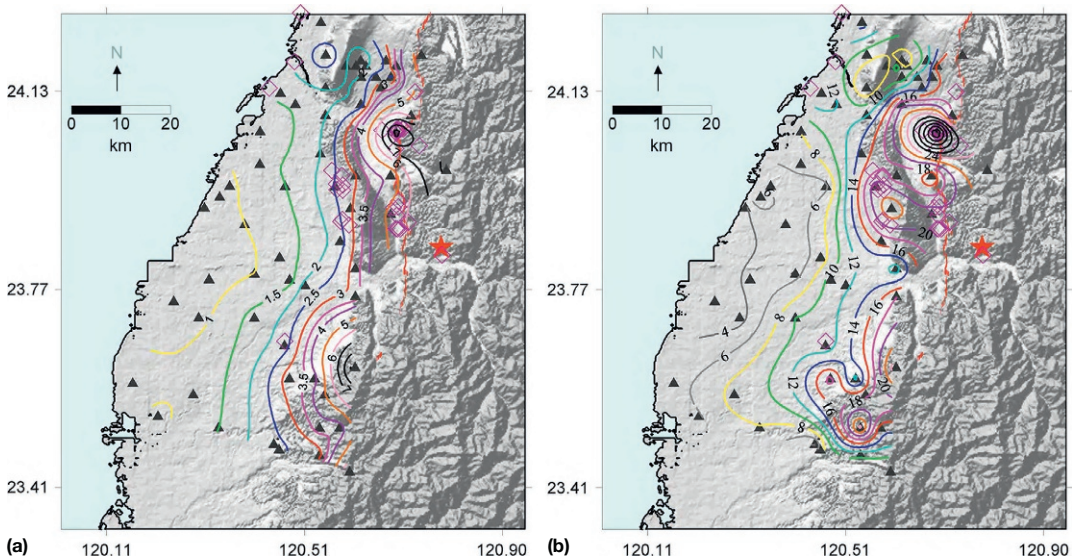


Figure 9 (a) Contours of the horizontal peak ground acceleration (PGA) during the Chi-Chi earthquake together with liquefaction sites in open diamonds. (b) Contours of the spectral acceleration at 1 Hz. The correlation of the occurrence of liquefaction with acceleration is better at 1 Hz than for the peak acceleration. See Wang et al. (2003) for a more detailed analysis. Modified from Wang C-Y, Dreger DS, Wang C-H, Mayeri D, and Berryman JG (2003) Field relations among coseismic ground motion, water-level change and liquefaction for the 1999 Chi-Chi (Mw = 7.5) earthquake, Taiwan. *Geophysical Research Letters* 17: http://dx.doi.org/10.1029/2003GL017601.

A permanent change in some aspect of pore structure is required, but the process by which small periodic strains change permeability or storage properties is unclear. Hydrogeochemical (e.g., Wang et al., 2004d) and temperature (e.g., Mogi et al., 1989; Wang et al., 2012; Wang et al., 2013) changes that accompany water-level changes confirm that pore-space connectivities change. One mechanism that has been proposed (to explain the particular postseismic change in water level shown in **Figure 4**) is that oscillatory flow back and forth in fractures caused by cyclic strain removes 'barriers' of fracture-blocking deposits, which then increases permeability and affects the final distribution of pore pressure (Brodsky et al., 2003). Elkhoury et al. (2006) confirmed that distant earthquakes can indeed change permeability. Earthquake-enhanced permeability does not seem to persist indefinitely. Manga et al. (2012) compiled observations of the timescale over which enhanced permeability returns to the preearthquake values. Timescales vary from as little as a few minutes (Geballe et al., 2011) to many years (e.g., Cappa et al., 2009; Kitagawa et al., 2007).

Other proposed, but also unverified, mechanisms include pore pressure increases caused by a mechanism 'akin to liquefaction' (Roeloffs, 1998), shaking-induced dilatancy (Bower and Heaton, 1978), or increasing pore pressure through seismically induced growth of bubbles (e.g., Linde et al., 1994). A feature of at least some far-field responses is that the magnitude of water-level changes scales with the amplitude of seismic waves (Sil and Freymueller, 2006).

One common feature of wells that exhibit far-field postseismic responses is that the sign of the water-level changes at a given well appears to be consistent among multiple earthquakes (Itaba and Koizumi, 2007; Matsumoto, 1992; Roeloffs, 1998; Sil and Freymueller, 2006). Any proposed mechanism should explain this feature, and the barrier explanation can only do so if the location of barriers is always located either up-gradient or down-gradient of the well. Wang and Chia (2008) suggested that this hypothesis makes a testable prediction that, if a sufficiently large number of observations are available, there should be a statistically random occurrence in both the sign and the magnitude of the water-level changes at a given epicentral distance. Using both a global set of water-level data and local data from Taiwan, Wang and Chia (2008) showed that there indeed appears to be such a statistically random occurrence. Far-field shaking-induced dilatancy (e.g., Bower and Heaton, 1978) or a process analogous to liquefaction (Roeloffs, 1998) would always cause the response to have the same sign, but neither can be universal mechanisms because the sign of water-level changes varies from well to well.

Observations from wells that have responded to multiple earthquakes indicate that wells may have an enhanced sensitivity over a specific range of frequencies (Roeloffs, 1998). Triggered earthquakes (Brodsky and Prejean, 2005), nonvolcanic tremor (Guilhem et al., 2010), liquefaction in the field (Holzer and Youd, 2007; Wong and Wang, 2007), and changes in eruption rate at mud volcanoes (Rudolph and Manga, 2012) all show increased sensitivity to long-period waves. Observations of, and models that account for, the frequency-dependent response may be useful for distinguishing between processes.

4.12.3.1.5 Summary

To generalize, on the basis of the reported changes in water level in the near and intermediate field, it appears that in general, deep wells in sound rock display a poroelastic response that mimics coseismic strain (in both sign and magnitude), whereas shallow wells and wells in poorly consolidated materials tend to show larger changes and increases in water level.

4.12.3.2 Streamflow

One of the most spectacular surface hydrologic responses to earthquakes is large changes in streamflow. **Figure 10** shows two typical examples. The increased discharge is persistent, at least until rainfall obscures the changes caused by the earthquake. The peak discharge can occur from within a day to as much as several weeks after the earthquake. Increased streamflow occurs in the near and intermediate field (Manga, 2001). The total excess discharge (i.e., the total volume of water released in excess of that expected in the absence of the earthquake) can be large: 0.7 km^3 for the M7.5 Chi-Chi earthquake (Wang et al., 2004b) and 0.5 km^3 after the M7.5 Hebgen Lake earthquake (Muir-Wood and King, 1993). Because streamflow responds to precipitation in the drainage basin, earthquake-induced changes are best recorded and studied during the dry season or during dry periods when there is little or no precipitation.

Explanations for changes in streamflow can be divided into six categories: expulsion of deep crustal fluids resulting from coseismic elastic strain (e.g., Muir-Wood and King, 1993), changes in near-surface permeability (Briggs, 1991; Rojstaczer

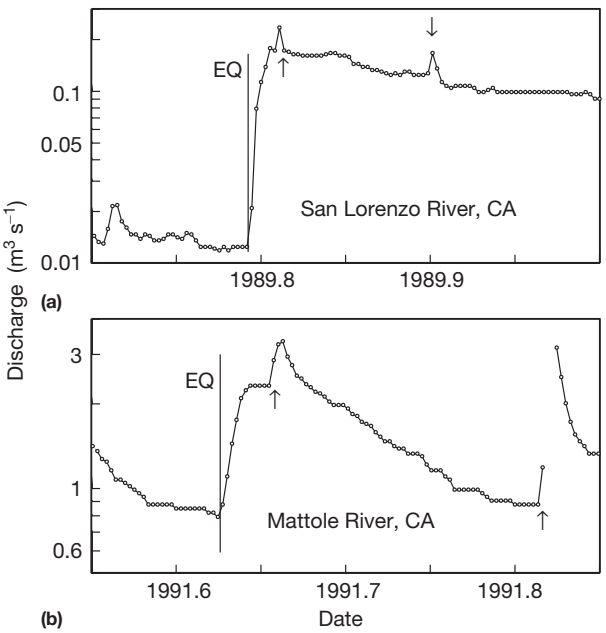


Figure 10 Hydrographs of (a) San Lorenzo River, CA, and (b) Mattole River, CA, showing postseismic response to the 1989 M7.3 Loma Prieta earthquake and 1991 M6 Honeydew earthquake, respectively. The vertical line labeled EQ indicates the time of the earthquake. Small arrows show the time of rainfall events. Figure made with US Geological Survey stream gauge data.

and Wolf, 1992; Rojstaczer et al., 1995; Sato et al., 2000; Tokunaga, 1999), consolidation or even liquefaction of near-surface deposits (Manga, 2001; Manga et al., 2003; Montgomery et al., 2003), release of water from the unsaturated zone (Manga and Rowland, 2009; Mohr et al., 2012), rupturing of subsurface reservoirs (Wang et al., 2004c), and the release of water trapped in fault zones. The differences between these different explanations are nontrivial because of their implications for the magnitude of crustal permeability, its evolution, and the nature of groundwater flow paths. We thus summarize the basis and some problems with each explanation and then consider the streamflow response to the 1999 Chi-Chi earthquake as an illustrative example:

1. Coseismic elastic strain

   Muir-Wood and King (1993) applied the coseismic elastic strain model, proposed by Wakita (1975) to explain coseismic groundwater-level changes, to explain the increased stream discharge after some large earthquakes, including the 1959 M7.5 Hebgen Lake earthquake and the 1983 M7.3 Borah Peak earthquake. Muir-Wood and King (1993) suggested that saturated microcracks in rocks were opened and closed in response to the stress change during an earthquake, causing pore volume to either increase or decrease, resulting in a decrease or increase in the groundwater discharge into streams.

   Following the 1989 Loma Prieta earthquake, however, chemical analysis of the stream water showed that the extra water that appeared in the streams following the earthquake had a shallow, rather than deep, origin (Rojstaczer and Wolf, 1992; Rojstaczer et al., 1995). Furthermore, Rojstaczer and colleagues showed that to account for the extra water in the increased streamflow by coseismic elastic strain, a very large volume of the crust must be involved in the expulsion of groundwater, which in turn would require an unreasonably high permeability for the crust. In addition, one important constraint on models is that streamflow generally increases, and the few streams with decreased discharge show very small changes.

   Manga et al. (2003) and Manga and Rowland (2009) reported that discharge increased at streams and springs in California after several earthquakes, irrespective of whether the coseismic strain in the basin was contraction or dilatation. For this reason, the dynamic strain caused by the earthquake, rather than the coseismic static strain, must be responsible for the observed changes in discharge, at least in the studied streams.

2. Enhanced permeability

   Briggs (1991) and Rojstaczer and Wolf (1992) and also Rojstaczer et al. (1995) proposed a model of enhanced permeability of the shallow crust resulting from seismically induced fractures and microfractures to explain the increased stream discharge in the nearby basins and the changes in the ionic concentration of stream water following the 1989 Loma Prieta earthquake in California. A shallow origin of the excess water is supported by a decrease in water temperature after discharge increased. Similar models were applied to the 1995 Kobe earthquake in Japan (Sato et al., 2000; Tokunaga, 1999) and the 2009 L'Aquila earthquake in Italy (Amoruso et al., 2011) to explain the observed hydrologic changes. Permeability enhancement was also invoked to explain increased electrical conductivity of water discharged after an earthquake (Charmoille et al., 2005). Elkhoury et al. (2006) found that earthquakes in Southern California caused phase shifts in the water-level response to tidal strain and interpreted these to be due to permeability increased by seismic waves. The increases in permeability are temporary; biogeochemical processes may act to reseal or block any newly created, fresh fractures.

   Manga (2001) found, however, that the rate of baseflow recession (i.e., the slow decrease in discharge during periods without precipitation) was unchanged by earthquakes that caused increases in streamflow. Assuming groundwater discharge Q to the stream is governed by Darcy's eqn [5], baseflow recession (a long time after recharge) is given by

$$\frac{dQ}{dt} \propto \exp[-aDt] \qquad [13]$$

   where a is a constant that characterizes the geometry of the drainage basin, D is the hydraulic diffusivity, and t is time. Although this is a great simplification of the complex subsurface flow paths of water to streams, such models in general provide an excellent empirical fit to discharge records. The slope of the hydrograph in **Figure 11(a)** is proportional to aD, and **Figure 11(b)** shows that the recession constant does not change following the earthquake. Given that the basin geometry and flow pathways are unlikely to have been reorganized by the earthquake, an unchanged baseflow recession implies that aquifer permeability did not change after the earthquake, even over the time period of increased discharge. This conclusion was substantiated by later studies in other areas (Montgomery et al., 2003) and during other earthquakes (Manga et al., 2003).

3. Coseismic consolidation and liquefaction

   Consolidation of loose materials is one way to increase pore pressures. Indeed, liquefaction occurred in many of the areas where streamflow increased (Montgomery and Manga, 2003). Consequently, Manga (2001) and also Manga et al. (2003) suggested that coseismic liquefaction of loose sediments on floodplains may provide the water for the increases in stream discharge following earthquakes.

   Figure 12 shows the relationship between earthquake magnitude and distance between the epicenter and the center of the gauged basin for streams that responded to earthquakes. Also shown for reference is the maximum distance over which liquefaction has been reported (from **Figure 2**). That the limit for both responses is similar does not imply that liquefaction causes streamflow to increase. Instead, the correlation simply means that dynamic strains sufficient to induce liquefaction under optimal conditions are also sufficient for streamflow to change.

   Wang et al. (2004b), however, used the extensive network of stream gauges in Taiwan to identify the source of excess discharge in streams. Excess discharge did not originate on the unconsolidated alluvial fan where there was, in fact, widespread liquefaction. Rather, the excess discharge originated in the mountains where there is little alluvial material and groundwater originated in bedrock.

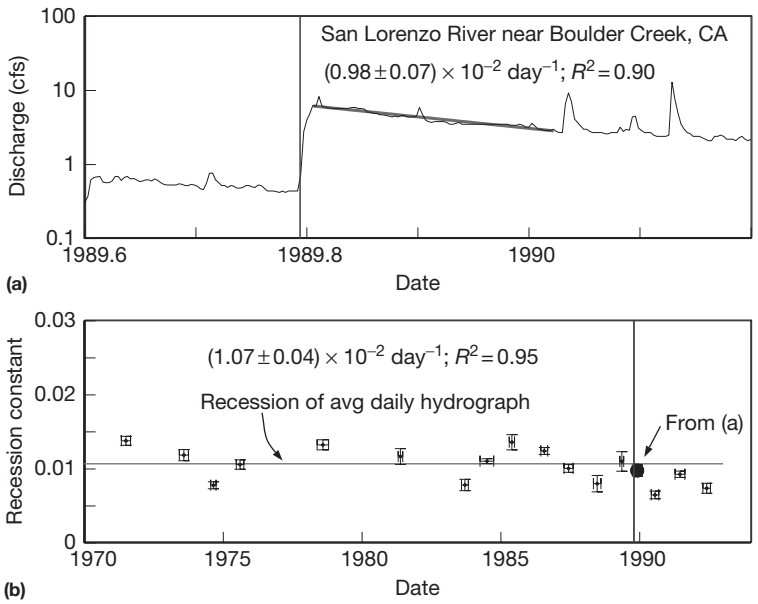


Figure 11 (a) Hydrograph of the San Lorenzo River, CA, showing postseismic response to the 1989 M7.3 Loma Prieta earthquake. The vertical line labeled EQ indicates the time of the earthquake. The postseismic period of baseflow recession is shown by the bold sloping line. (b) The baseflow recession constant for periods of baseflow before and after the earthquake showing that even though discharge increased by an order of magnitude, there was no significant change in baseflow recession. Figure made with US Geological Survey stream gauge data. Modified from Manga M (2001) Origin of postseismic streamflow changes inferred from baseflow recession and magnitude-distance relations. *Geophysical Research Letters* 28: 2133–2136.

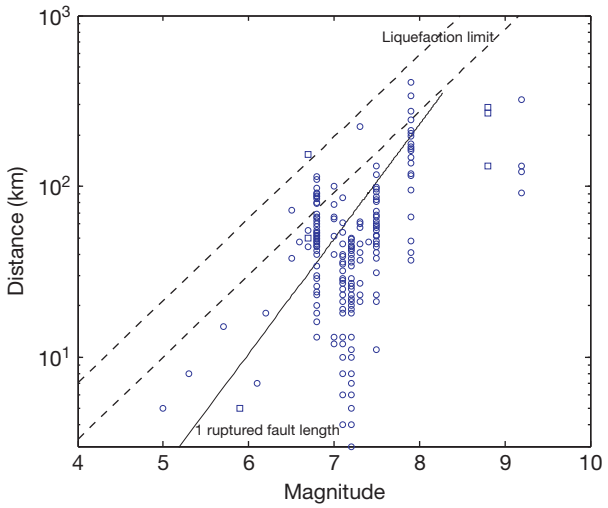


Figure 12 Relationship between earthquake magnitude and distance from the epicenter of streams that exhibited clear and persistent increases in discharge. Data shown with circles are from the compilation in Wang and Manga (2010a) with additional data shown with squares for the M8.8 Maule earthquake of 2010 (Mohr et al., 2012), the Tasso and Tanagro river responses to the M6.7 Irpinia, Italy, earthquake of 1930 (Esposito et al., 2009), and spring responses to the M5.9 Draney Peak, the United States, earthquake of 1994 (Schuster and Murphy, 1996). The dashed lines bound the maximum distance for liquefaction and the solid line is 1 ruptured fault length (see **Figure 2**).

In addition, after the 2001 M6.8 Nisqually, Washington, earthquake, Montgomery et al. (2003) did not find a significant field association between liquefaction occurrence and streamflow changes.

4. Water release from the unsaturated zone

   In some cases, excess discharge appears to originate from the unsaturated zone. Mohr et al. (2012) found that evapotranspiration and discharge both increased in small Chilean catchments after the M8.8 2010 Maule earthquake, suggesting that water drained from the unsaturated zone to the water table. Manga and Rowland (2009) noted that chloride concentrations and stable O and H isotopes in stream water after an earthquake were consistent with water that underwent significant evapotranspiration, a process that occurs in and near the unsaturated zone.

5. Fault valves and ruptured subsurface reservoirs

   In convergent tectonic regions, large volumes of pore water may be locked in subducted sediments (Townend, 1997). Sealing may be enacted partly by the presence of low-permeability mud and partly by the prevailing compressional stresses in such tectonic settings (Sibson and Rowland, 2003). Earthquakes may rupture the seals and allow pressurized pore water to erupt to the surface and recharge streams. This process may explain time variations in submarine fluid discharge at convergent margins (Carson and Screaton, 1998). Episodes of high discharge are correlated with seismic activity having features similar to tremor but are not correlated with large regional earthquakes (Brown et al., 2005).

   Following the M6.5 San Simeon, CA, earthquake on 22 December 2003, a different type of streamflow increase was revealed in the central Coast Ranges, occurring within 15 min of the earthquake and lasting about an hour. At the same time, several new hot springs appeared across a dry riverbed, a short distance upstream of the stream gauge that

registered the abrupt increase in streamflow. These observations suggest that the occurrence of streamflow increase and the appearance of hot springs were causally related and may be the result of earthquake-induced rupture of the seal of a hydrothermal reservoir. Recession analysis of the stream gauge data shows $D/L^2 \sim 1000\,s^{-1}$, where L is the depth of the reservoir, implying a small L or large D or both. Assuming a one-dimensional model for flow between the ruptured seal and the stream, Wang et al. (2004c) estimated that the extra discharge after this earthquake was $\sim 10^3\,m^3$. Abrupt changes in hot spring discharge are known to occur in the Long Valley, CA, hydrothermal area (Sorey and Clark, 1981), suggesting that this type of hydrologic response may be characteristic of hydrothermal areas.

Rupturing of permeability barriers may occur at depths of many kilometers. Husen and Kissling (2001) suggested that postseismic changes in the ratio of P-wave and S-wave velocities above the subducting Nazca Plate reflect fluid migration into the overlying plate following the rupture of permeability barriers. Disruption of low-permeability barriers by seismic waves generated $>10^4$ km away has been proposed (Doan and Cornet, 2007).

To illustrate some approaches that are used to test these explanations, we again consider the records from the 1999 M7.5 Chi-Chi earthquake in central Taiwan. In central Taiwan, 17 stream gauges are located on three stream systems (**Figure 13(a)**) that showed different responses to the 1999 Chi-Chi earthquake (Wang et al., 2004b). The local rainy season was over well before the earthquake. These records have made it possible to test the different hypotheses listed earlier. Among the three streams, two (Choshui Stream and the Wushi Stream) have extensive tributaries in the mountains (**Figure 13(a)**), but the third (Peikang Stream) originates on the sloping side of the Choshui fan, with no tributaries in the mountains. Streams in the mountains all showed large postseismic streamflow increases. On the alluvial fan, both the Choshui Stream and the Wushi Stream, which have tributaries in the mountains, showed large increases in streamflow. The amount of increase in streamflow in the proximal area of the Choshui alluvial fan was about the same as that in the distal area of the fan, indicating that most of the excess water originated in the mountains and that the contribution from the sediments in the fan was insignificant. By contrast, the Peikang Stream system, which does not have tributaries in the mountainous area, did not show any noticeable postseismic increase in streamflow. Thus, it may be concluded that the excess water in the streams after the Chi-Chi earthquake requires an earthquake-induced discharge from the mountains. Any coseismic consolidation and liquefaction of sediments on the floodplain (alluvial fan) were not substantial enough to cause noticeable postseismic increase in streamflow.

Figure 13(b) shows the discharge of a stream in the mountains before and after the Chi-Chi earthquake, together with the precipitation record from a nearby station. In addition to a coseismic increase in streamflow and a slow postseismic return to normal, the figure shows that the baseflow recession rates both before and after the earthquake are identical. This observation is of interest because it suggests that the Chi-Chi earthquake did not cause significant changes in the permeability of the aquifer that fed the stream, as Manga (2001) first noticed in

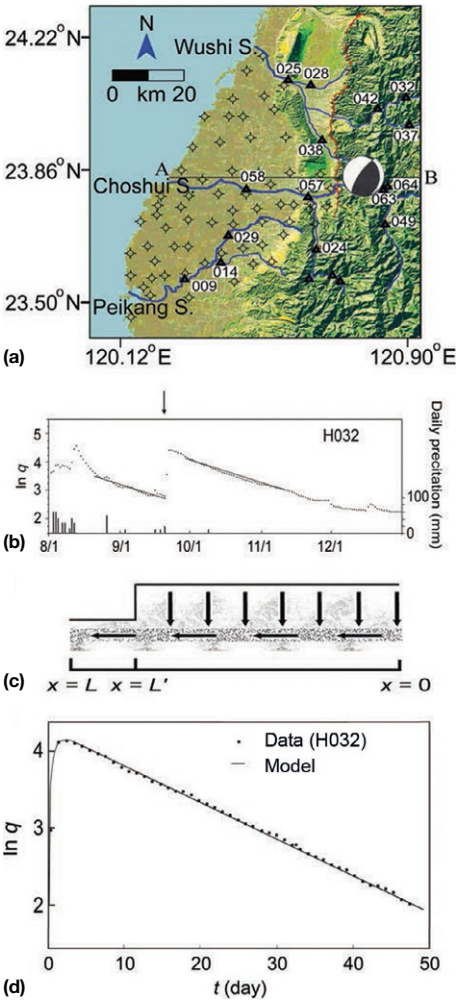


Figure 13 (a) Map showing the three river systems in central Taiwan and the locations of the stream gauges (triangles). Note that both the Wushi River and the Choshui River have extensive tributaries in the mountains, but the Peikang River does not. Hydrologic monitoring wells are marked by circles overlying crosses. (b) Logarithm of stream discharge as recorded by a gauge in central Taiwan as a function of time before and after the Chi-Chi earthquake, together with precipitation record at a nearby station. Note that the recession curve before and after the earthquake (indicated by arrow) has the same slope. (c) The conceptual model showing increased recharge of aquifer by groundwater from mountains due to enhanced vertical permeability during and after the earthquake. (d) Comparison of excess discharge as a function of time with model result (thin curve). Excess discharge at the time of earthquake ($t = 0$) is taken to be zero and therefore does not appear on the diagram.

his study of the coseismic responses of streams in the United States. However, the excess water in streams in central Taiwan after the Chi-Chi earthquake requires extra discharge from the mountains that can appear only through enhanced permeability during the earthquake. This apparent contradiction may be resolved by recognizing that the earthquake-induced change in permeability may be anisotropic. The foothills of the Taiwan mountains consist of alternating beds of sandstone and shale. Thus, before the earthquake, the vertical permeability would be controlled by the impervious shales and groundwater flow

would be mostly subhorizontal. During the earthquake, subvertical fractures breached the impervious shales and enhanced vertical permeability, allowing rapid downward draining of water to recharge underlying aquifers. The horizontal conductivity of the aquifers, which was high before the earthquake, was essentially unaffected. Field surveys after the Chi-Chi earthquake revealed numerous subvertical tensile cracks in the hanging wall of the thrust fault (Angelier et al., 2000; Lee et al., 2000, 2002). Many wells in the foothills above the thrust fault showed a significant drop in water level (Chia et al., 2001; Lin, 2000; Wang et al., 2001; Yan, 2001), and a sudden downpour occurred in a tunnel beneath the foothills right after the earthquake. Moreover, water temperatures decreased in wells located near the upper rim of the alluvial fan, as expected if vertical permeability and vertical groundwater flow increased (Wang et al., 2012). These observations are consistent with the model of enhanced vertical permeability during the Chi-Chi earthquake.

A one-dimensional leaky aquifer model (**Figure 13(c)**) may be used to quantify this conceptual model. Flow is governed by eqn [7], and the recharge to the aquifer resulting from enhanced vertical permeability is treated as a source. Even though this model is highly simplified, several studies (e.g., Brodsky et al., 2003; Manga, 2001; Manga et al., 2003; Roeloffs, 1998) have demonstrated that it is useful for characterizing the first-order response of hydrologic systems to earthquakes.

We treat the recharge of the leaky aquifer as coseismic with constant recharge for $0 < x \le L'$ (length of aquifer beneath the mountain with enhanced permeability) and no recharge for $L' > x > L$ (length of aquifer beneath the floodplain). If we further assume a boundary condition $h = 0$ at $x = L$ and a no-flow condition at $x = 0$, the excess discharge to the stream at time t is given by (e.g., Wang et al., 2004a)

$$Q_{ex} = \frac{2DVQ_0}{L^2(L'/L)} \sum_{n=1}^{\infty} (-1)^{n-1} \sin\left[\frac{(2r-1)\pi L'}{2L}\right] \exp\left[-\frac{(2r-1)^2\pi^2 D}{4L^2} t\right] \quad [14]$$

where L is the total length of the aquifer and V is the volume of the aquifer between $x = 0$ and $x = L'$ (**Figure 13(a)**); thus, VQ_0 is the total excess water recharging the aquifer. For large times, eqn [14] results in an exponential decrease in discharge, the so-called baseflow recession given by eqn [13]. **Figure 13(d)** compares the data for stream gauge 032 (adjusted to $Q_{ex} = 0$ just before the earthquake) with the modeled excess discharge Q_{ex} determined by eqn [14]. An excellent fit is obtained with $D/L^2 = 2.4 \times 10^{-7}\,s^{-1}$ (from recession analysis) and $L'/L = 0.8$. The latter value is consistent with the fact that gauge 032 is located in the mountains, where the floodplain is small in comparison with the total aquifer length. The amount of excess water VQ_0 may be determined from the value of DVQ_0/L^2 from model fitting. By summing the excess discharges in the Choshui Stream and in the Wushi Stream, we obtain a total amount of 0.7–0.8 km^3 for the excess water released from the mountains after the Chi-Chi earthquake.

For the example shown in **Figure 13**, the peak postseismic discharge occurs within a day of the earthquake. **Figure 14** shows an example in which the peak discharge is reached 9–10 days later, though discharge begins to increase coseismically. This example is for Sespe Creek, CA, responding to the 1952 M7.5 Kern County earthquake located 63 km away from the center of the drainage basin. Again, the model for excess discharge given by eqn [14], shown by the solid curve in **Figure 14**, fits the observed postseismic discharge very well (baseflow described by eqn [13] has been added back to the calculated excess discharge).

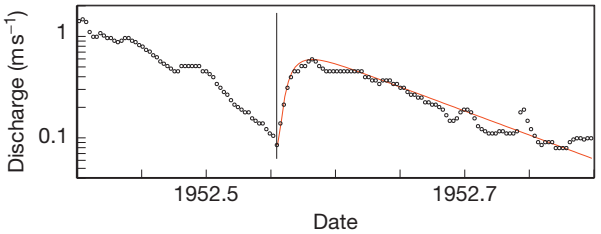


Figure 14 Response of Sespe Creek, CA, to the 1952 M7.5 Kern County earthquake. Daily discharge measurements collected and provided by the US Geological Survey are shown with circles. Curve is the solution to eqn [14] for the excess flow with $L'/L = 0.4$ added to the baseflow, eqn [13], to recover the entire hydrograph. Vertical line shows the time of the earthquake. There was no precipitation during the time interval shown in this graph.

The different hypotheses listed in this section imply different crustal processes and different water–rock interactions during an earthquake cycle. In most instances, the hydrologic models are underconstrained. A reasonable approach is to test the different hypotheses against cases (such as Chi-Chi) in which abundant and accurate data are available. We note that a single explanation need not apply for all cases of increased streamflow so that identifying when and where different mechanisms dominate is important.

4.12.3.3 Mud Volcanoes

Mud volcanoes range from small, centimeter-size structures to large edifices up to few hundred meters high and several kilometers across. Mud volcanoes erupt predominantly water and fine sediment. They occur in regions where high sedimentation rates and fine-grained materials allow high pore pressures to develop (Kopf, 2002). Large mud volcanoes erupt from depths of several hundred meters to more than a couple kilometers.

Mud volcanoes also erupt in response to earthquakes (e.g., Panahi, 2005) and thus provide another probe of earthquake–hydrology interactions. Unfortunately, though, the number and quality of reports prevent a rigorous statistical analysis of the correlation (Bonini, 2009; Manga and Brodsky, 2006; Manga et al., 2009; Mellors et al., 2007).

A necessary condition to create mud volcanoes is the liquefaction or fluidization of erupted materials so that the erupted materials lose strength and can behave in a liquid-like manner (Pralle et al., 2003). Unconsolidated sediment can be liquefied and fluidized through mineral dehydration, gas expansion, tectonic stresses, inflow of externally derived fluids, and even ocean waves (Maltman and Bolton, 2003).

Figure 15 shows that the occurrence of triggered mud volcanism and changes in eruption rate falls within the liquefaction limit for shallow (upper few to tens of meters) liquefaction. The data shown in **Figure 15** are compiled and

tabulated in Wang and Manga (2010a) with additional data from Rudolph and Manga (2010, 2012), Manga and Bonini (2012), and Tsunogai et al. (2012).

Liquefaction is usually thought to be a shallow phenomenon confined to the upper few to tens of meters. This is because the fluid pressure needed to reach lithostatic pressure usually increases with depth. However, sedimentary basins with high sedimentation rates, fine sediments (with low permeability), and lateral compression often have high fluid pressures. Indeed, it is these settings in which mud volcanoes seem to occur (Milkov, 2000). Seismically triggered liquefaction may not necessarily be only a shallow phenomenon.

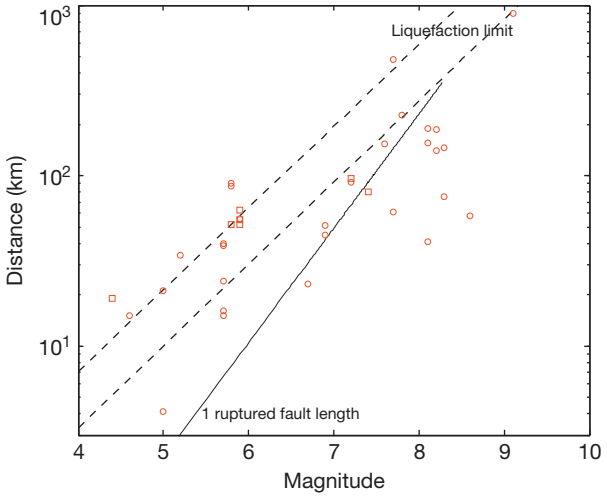


Figure 15 Relationship between earthquake magnitude, distance from the epicenter, and eruption of large mud volcanoes that originate from depths greater than hundreds of meters and erupted within a few days of the earthquake. Circles show data compiled and tabulated in Wang and Manga (2010a); squares show additional data from the Salton Sea, CA (Rudolph and Manga, 2010, 2012), the Northern Apennines, Italy (Manga and Bonini, 2012), and submarine mud volcanoes (Tsunogai et al., 2012). The dashed lines bound the maximum distance for liquefaction and the solid line is 1 ruptured fault length (see **Figure 2**).

Processes other than liquefaction have been invoked to explain the response of mud volcanoes and mud eruptions to earthquakes. Mazzini et al. (2009) appealed to reactivation of faults to enhance or create pathways to the surface that release overpressure. Rudolph and Manga (2012) showed that mud volcanoes are more sensitive to long-period waves of a given amplitude than short-period waves with the same amplitude. This supports mechanisms in which time-varying flows produced by seismic waves mobilize trapped bubbles (Beresnev et al., 2011) because bubble mobilization is more effective at low frequencies (Beresnev, 2006). Mobilizing or dislodging trapped bubbles has the effect of increasing permeability.

4.12.3.4 Geysers

Geysers change eruption frequency following distant earthquakes that generate coseismic static strains smaller than 10^{-7} or dynamic strains less than 10^{-6} (e.g., Hutchinson, 1985; Silver and Valette-Silver, 1992). Geysers, despite requiring very special thermal conditions and hydrogeologic properties to occur (Steinberg et al., 1982a), thus provide another opportunity to understand how earthquakes can influence hydrogeologic processes and properties.

Figure 16 shows the response of Daisy Geyser in Yellowstone National Park to the 2002 M7.9 Denali earthquake

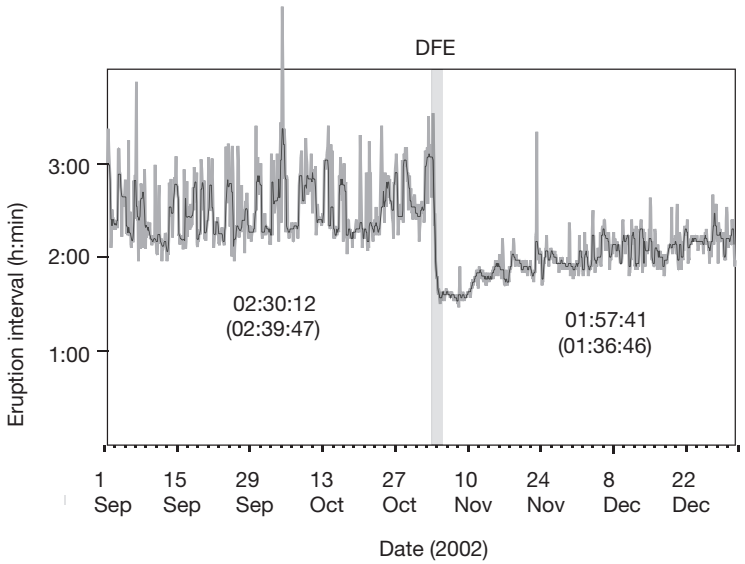


Figure 16 Eruption interval (time between two successive eruptions) in hour:minute format as a function of time at Daisy Geyser, Yellowstone National Park. Thin gray line shows raw data, that is, the actual measured eruption intervals. The black line is smoothed data obtained by averaging over several measurements. Median eruption intervals are also shown in hour/minute/second format. Median eruption intervals are computed over a several week period; intervals in parentheses are based on a few days before and after the earthquake. Time of the 2002 M 7.9 Denali earthquake is shown by the vertical line. Figure provided by Stephan Husen and made of data presented in Husen S, Taylor R, Smith RB, and Healser H (2004) Changes in geyser eruption behavior and remotely triggered seismicity in Yellowstone National Park produced by the 2002 M7.9 Denali fault earthquake, Alaska. *Geology* 32: 537–540.

located more than 3000 km away (Husen et al., 2004). The eruption interval decreases by about a factor of 2 and then slowly increases to the preearthquake frequency over a period of a few months. Among the many geysers at Yellowstone, the eruption frequency of some increased, whereas for others, it decreased.

Although geysers often respond to earthquakes, their response to barometric pressure changes (e.g., White, 1967), solid earth tides (e.g., Rinehart, 1972), and hypothetical preseismic strains (Silver and Valette-Silver, 1992) is weaker. The apparent correlation of eruption interval and nonseismic strains is, in fact, controversial. Rojstaczer et al. (2003) reanalyzed eruption records at Yellowstone and found that geysers were insensitive to earth tides and diurnal barometric pressure changes and thus concluded that geysers were not sensitive to strains less than about $10^{-8} - 10^{-7}$. One of these records was from Daisy Geyser (**Figure 16**). Given that Daisy Geyser responded to earthquakes that produced static strains as small as the amplitude of tidal strains, we can reasonably conclude that it is the larger amplitude dynamic strain, rather than static strain, that causes the observed responses (e.g., Husen et al., 2004).

The mechanism by which the eruption interval changes is not known. Steinberg et al. (1982b) noted that one mechanism for initiating a geyser eruption is superheating the water. Using a lab model of a geyser, Steinberg et al. (1982c) showed that vibrations can reduce the degree of superheating and can thus increase the eruption frequency. It is worth noting that while earthquakes can decrease the interval between eruptions, sometimes, the frequency of eruptions decreases. Ingebritsen and Rojstaczer (1993) proposed instead that much of the observed temporal variability in eruption frequency can be explained by changes in matrix permeability, which in turn governs the recharge of the conduit. Geyser frequency is also sensitive to the (often complex) conduit geometry (Hutchinson et al., 1997) and hence permeability of the geyser conduit. However, because the conduit is already much more permeable, small strains within the conduit itself are unlikely to have a significant effect on geyser eruptions (Ingebritsen et al., 2006). As a consequence, the observed temporal variability in eruption frequency is probably caused by changes in the permeability of the matrix surrounding the main geyser conduit – it is through this matrix that the conduit is recharged between eruptions.

The high sensitivity and long-lasting response of geysers to small seismic strains require reopening, unblocking, or creating fractures to induce large enough permeability changes to influence eruption frequency. Geysers thus provide evidence that dynamic strains are able to create permanent changes in permeability from small dynamic strains and at great distances from the earthquake.

In summary, the response of geysers to distant earthquakes is most easily explained by changes in permeability, and the sensitivity of geysers to earthquakes compared with earth tides and changes in barometric pressure indicates that dynamic strains cause the response.

4.12.4 Feedback Between Earthquakes and Hydrology

The presence of water in the subsurface, and changes in the amount of water on the surface or within the subsurface, can influence the occurrence of earthquakes, as summarized in **Figure 17**. There are three basic ways water can influence seismicity.

First, changes in loading of the surface can increase deviatoric stresses. The most clear examples follow the impoundment of water by reservoirs (e.g., Gupta, 1992; Simpson et al., 1988). Natural examples are more difficult to identify, probably in part because the magnitude of the surface load is much smaller; proposed examples include snow loading (Heki, 2003) and loading from ocean tides (Cochran et al., 2004) or other seasonal processes (Wolf et al., 1997).

Second, changes in fluid pressure p reduce the effective stress

$$\sigma_{ij}^{\text{eff}} = \sigma_{ij} - \alpha p \delta_{ij} \qquad [15]$$

where α is again the Biot–Willis coefficient. The classic example of seismicity induced by changes in pore pressure is earthquakes caused by fluid injection (e.g., Raleigh et al., 1976). Some cases of reservoir-induced seismicity, in particular those

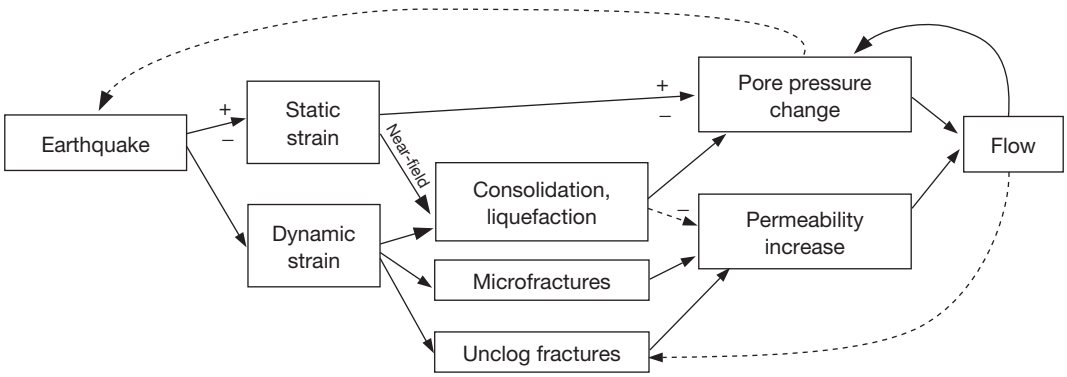


Figure 17 Schematic illustration of the relationship between earthquakes and hydrologic responses. + and – signs indicate the sign of the response. Figure based on an illustration originally created by David Mays. See Manga M, Beresnev I, Brodsky EE, Elkhoury JE, Elsworth DE, Ingebritsen S, Mays DC, and Wang C-Y (2012) Changes in permeability by transient stresses: Field observations, experiments and mechanisms. *Reviews of Geophysics* 50: RG2004, http://dx.doi.org/10.1029/2011RG000382.

in which the seismicity lags changes in water level, have been attributed to changes in pore pressure p (e.g., do Nascimento et al., 2005; Ge et al., 2009; Talwani and Acree, 1985). Changes in pore pressure also diffuse, governed by eqn [6], so the telltale signature of earthquakes triggered by changes in pore pressure would be a migration of the locus of seismicity at a rate consistent with the (usually unknown) hydraulic diffusivity. Spatial migration of reservoir-induced seismicity (Talwani and Acree, 1985), seismicity induced by fluid injection (Parotidis and Shapiro, 2004; Shapiro et al., 2006; Tadokoro et al., 2000), and some aftershock sequences (e.g., Bosl and Nur, 2002; Hainzl, 2004; Miller et al., 2004; Nur and Booker, 1972) have been attributed to pore pressure diffusion. Pore pressure changes caused by natural loading at the surface (e.g., groundwater recharge) have also been invoked to trigger seismicity (e.g., Christiansen et al., 2005, 2007; Costain et al., 1987; Hainzl et al., 2006; Saar and Manga, 2003).

Third, fluid extraction, rather than fluid injection, can change stresses through poroelastic deformation (Segall, 1989). Earthquakes appear to have been induced by fluid extraction in gas fields (Segall et al., 1994) and oil fields (Gomberg and Wolf, 1999; Zoback and Zinke, 2002). A decrease in pore pressure might be expected to stabilize faults by increasing the effective stress. Poroelastic deformation caused by changes in fluid pressure can, however, increase the magnitude of deviatoric stresses away from the region of fluid extraction.

One general conclusion from all these studies of hydrologically triggered seismicity, regardless of the mechanism, is that very small stress changes, typically 0.01–1 MPa, appear to trigger the earthquakes. Such small stress changes, however, are also similar to the stress changes thought to trigger earthquakes through nonhydrologic means (e.g., Harris, 1998; Stein, 1999).

If earthquakes cause changes in hydrogeologic properties and changes in fluid pressure promote seismicity, then hydrogeologic and seismological processes are coupled. Triggered seismicity (see review by Hill and Prejean; **Chapter 4.11**) is one possible example of this coupling. Interaction is promoted if the state of stress is close to failure so that the small changes in stress associated with either natural hydrologic processes or hydrologic responses to earthquakes can in turn influence seismicity. At least in tectonically active areas, many faults do appear to be critically stressed (Zoback et al., 1987), consistent with earthquakes being triggered by small stress changes.

Rojstaczer et al. (2008) suggested that one manifestation of this interaction is the value of the mean large-scale permeability of the crust – it should be of a size to accommodate internal forcing (fluid generation by metamorphic and magmatic processes) and external (groundwater recharge and discharge) forcing. Diagenesis, in general, tends to seal fractures and fill pores, thus decreasing permeability. If groundwater recharge does not change with time, water levels and pore pressures will increase, promoting seismicity, which in turn generates new fractures and increases permeability. A mean crustal permeability of 10^{-14}m^2 (Manning and Ingebritsen, 1999) is consistent with the present mean rate of groundwater recharge and hydraulic head gradients driving basin-wide groundwater flow. Rojstaczer et al. (2008) proposed that this balance is reached by a feedback between hydrologic processes and seismicity. A similar balance must occur in the lower crust, though the source of water is from metamorphic reactions rather than of meteorologic origin. The mean permeability obtained from such a balance, however, is a temporal and spatial average. Instead, evidence of short-lived and locally high permeabilities is preserved in the spatial distribution of mineral deposits in ancient fault systems (e.g., Micklethwaite and Cox, 2004) and transient, localized high temperatures in the lower crust (Camacho et al., 2005).

4.12.5 Hydrologic Precursors

As noted in **Section 4.12.2.1**, failure of brittle rocks under deviatoric stress is usually preceded by pervasive microcrack formation. In the subsurface, this process (i.e., microfracturing) would greatly increase the surface area of the affected rock in contact with groundwater and thus allow the release of gases and dissolved ions from the rock into the groundwater, changing its chemical composition. Coalescence of microcracks into larger fractures may connect hydraulically isolated aquifers, causing both changes in the hydraulic heads and mixing of groundwater with initially distinct hydrogeochemistry. These processes may further cause changes in the electrical conductivity of the rocks and thus of the crust. Scenarios of this kind have led to the not unreasonable expectation that hydrologic, hydrogeochemical, and related geophysical precursors may appear before the occurrence of large earthquakes.

During the 1960s and 1970s, several groups reported precursory changes in the crustal seismic velocity ratio (V_p/V_s) before some earthquakes in the then Soviet Union (Savarensky, 1968; Semenov, 1969), in New York (Aggarwal et al., 1973), and in California (Whitcomb et al., 1973). Scholtz et al. (1973) developed a dilatancy model as a common physical basis for precursory phenomena. Coupling the failure processes of brittle rocks with a groundwater flow model, Nur (1974) proposed a dilatancy model to explain the sequence of precursory changes in V_p/V_s. Later seismic experiments designed to detect changes in both V_p and V_s, however, failed to reveal any precursory changes before many major and moderate earthquakes (e.g., McEvilly and Johnson, 1974). In 1988, in response to Bakun and Lindh's (1985) prediction that a large earthquake would occur near Parkfield before 1993, Park installed a network of electrodes across the fault zone near Parkfield to detect any precursory changes in the electrical resistivity of the crust (Park, 1997) – one of the major predictions of the dilatancy model (Scholtz et al., 1973). The predicted Parkfield earthquake did not occur; but a M6.0 earthquake occurred on 28 September 2004, near Parkfield, providing an excellent opportunity to test the telluric method for detecting any precursory changes in crustal resistivity before a large earthquake. Careful data processing was applied to the time series of the dipole fields before the earthquake, but no precursory changes were found (Park, 2005). These failures, among others, have called into question the general validity of the dilatancy model, and about earthquake precursors in general.

A concerted search for hydrologic and hydrochemical precursors to earthquakes has also been pursued, and many candidates have been reported in the literature. For example, a few days prior to the 1946 M8.3 Nankaido earthquake in Japan,

water levels in some wells reportedly fell by more than 1 m and some wells went dry (Linde and Sacks, 2002; Sato, 1982). Three days before the 1985 M6.1 Kettleman Hills, CA, earthquake, Roeloffs and Quilty (1997) found a gradual, anomalous rise in water level of 3 cm. This observation was included in the IASPEI Preliminary List of Significant Precursors (Wyss and Booth, 1997). Following the 1995 M7.2 Kobe earthquake, several papers reported precursory changes in the concentrations of radon, chlorine, and sulfate ions in groundwater (e.g., Igarashi et al., 1995; Tsunogai and Wakita, 1995) and in groundwater level (King et al., 1995). Changes in radon concentration are the most commonly reported and discussed hydrogeochemical precursor (e.g., Trique et al., 1999; Wakita et al., 1988) because the release of radon is especially sensitive to crustal strains.

Definitive and consistent evidence for hydrologic and hydrochemical precursors, however, has remained elusive (Bakun et al., 2005; Wang and Manga, 2010a). There are many difficulties: Most reported changes were not corrected for the fluctuations in temperature, barometric pressure, earth tides, and other environmental factors, so that some changes taken to be earthquake-related may in fact be 'noise' (e.g., Hartman and Levy, 2005); changes were recorded at some sites but often were not recorded at other nearby sites (e.g., Biagi et al., 2001); instrument failures and personnel/program changes often do not allow persistent and consistent monitoring over long periods of time (King et al., 2000) – a necessary condition for obtaining reliable precursory data; it is technically challenging to perform continuous monitoring of chemical parameters (dissolved solids or gases), which might be more sensitive than easier-to-measure physical parameters such as water level, discharge, or temperature. Distinguishing a precursor from a response to a previous earthquake (e.g., the 1946 Nankaido earthquake mentioned in the previous paragraph was preceded by the 1944 M8.2 Tonankai event) creates additional ambiguity. Indeed, care is needed to distinguish a precursor from a hydrologic response. For example, Plastino et al. (2010) documented uranium groundwater anomalies at four sites prior to and after the 6th April Mw 6.3 L'Aquila, Italy, earthquake. They documented four excursions in uranium concentration at two sampling sites in the 10 months prior to the earthquake and then no excursions in the 2 months after the earthquake. Plastino et al. (2010) proposed that these excursions reflect strain in a preparatory phase. This interpretation, however, does not explain which anomalies will be followed by an earthquake, and the absence of postevent excursions (if real) could simply reflect postseismic changes. The latter is expected given that there were postseismic changes in stream and spring discharge.

Notwithstanding many difficulties, progress has been made in the past decade. For example, intensive and continued observations of various kinds of potential precursory hydrologic and hydrochemical changes have been made in Japan during the past half century (Wakita, 1996), and records are routinely corrected to remove the noise introduced by fluctuations in temperature, barometric pressure, earth tides, and other factors (Igarashi and Wakita, 1995); methods for filtering and time series analysis have been developed and applied to time series of raw hydrochemical and water-level data to remove changes unrelated to earthquake processes (Chen et al., 2010; Kingsley et al., 2001); and relationships among various types of hydrologic and hydrochemical precursory signals are being probed, as these would provide a better foundation for identifying and understanding any precursors (Hartman and Levy, 2005). It is now clear that long time series of observations are required: Claesson et al. (2007) found that an extended time series of geochemical measurements reduced the statistical significance of possible hydrogeochemical precursors they had previously postulated (Claesson et al., 2004). Although we may still be far from achieving a genuine understanding of the underlying mechanisms of the various earthquake-related anomalies, significant efforts are under way.

Static strains of 10^{-8} can reasonably be expected to produce changes in water level of about 1 cm for optimal hydrogeologic properties. Even greater sensitivity to strain (though perhaps only dynamic strains) is implied by some of the very distant hydrologic responses seen at geysers and some very distant changes in water in wells (e.g., Brodsky et al., 2003). The strong sensitivity of hydrologic processes and properties to small strains is the primary basis for hope that hydrologic and hydrogeochemical monitoring may detect any hypothetical preearthquake strains. Recognizing precursors, and distinguishing between responses from previous earthquakes and precursors to new earthquakes, is a different matter and may continue to be problematic (Hartman and Levy, 2005). Although it is unclear whether any future documented hydrologic precursors could actually be used for earthquake prediction, their occurrence or absence may at least provide new insight into the physics of earthquakes and the tectonic processes that lead to earthquakes.

4.12.6 Concluding Remarks

Hydrologic responses to earthquakes provide constraints on hydrogeologic processes in regions that might otherwise be inaccessible, for example, fault zones or the deep subsurface. Measured responses may also provide information at spatial and temporal scales that are difficult to study with more conventional hydrogeologic measurements such as well tests. However, these novel features of hydrologic responses also mean that it can be difficult, and perhaps even impossible in some cases, to obtain the information needed to distinguish between competing models for hydrologic responses.

Explaining the hydrologic responses to earthquakes should in principle be simple because they reflect the strain caused by earthquakes. The great variety of hydrologic responses, however, highlights the complexity of deformation and structure of geologic materials and the interaction among processes. Over the last decade, there has thus been a trend toward developing quantitative physically based models to explain hydrogeologic responses to earthquakes with an emphasis on explaining phenomena that cannot be explained by linear poroelastic models alone.

We conclude by noting that advances in physically based models may not be sufficient for understanding the interactions between earthquakes and water. Although expensive and time-consuming, continued monitoring of wells, springs, and streams, ideally at high sampling rates and with

complementary data sets (e.g., chemistry, temperature, and pressure), provides the data needed to test models and may also lead to the discovery of new hydrologic phenomena.

Acknowledgments

We thank Alex Wong, Fu-Qiong Huang, Maria Brumm, Emily Brodsky, Hiroo Kanamori, Steve Ingebrisen, and three reviewers for comments and suggestions, though none of these individuals should be held responsible for views and biases in this chapter. We also thank Fu-Qiong Huang for providing Figure 2 in this chapter. This is supported by the National Science Foundation.

References

Aggarwal YP, Sykes LR, Armbruster J, and Sbar ML (1973) Premonitory changes in seismic velocities and prediction of earthquakes. *Nature* 241: 101–104.

Akita F and Matsumoto N (2004) Hydrological responses induced by the Tokachi-oki earthquake in 2003 at hot spring well in Hokkaido, Japan. *Geophysical Research Letters* 31. http://dx.doi.org/10.1029/2004GL020433.

Ambraseys NN (1988) Engineering seismology. *Earthquake Engineering and Structural Dynamics* 17: 1–105.

Amoruso A, Crescentini L, Petitta M, Rusi R, and Tallini M (2011) Impact of the 6 April 2009 L'Aquila earthquake on groundwater flow in the Gran Sasso carbonate aquifer, Central Italy. *Hydrological Processes* 25: 1754–1764.

Andrew DJ (2003) A fault constitutive relation accounting for thermal pressurization of pore fluid. *Journal of Geophysical Research* 107. http://dx.doi.org/10.1029/2002JB001942.

Angelier J, Chu HT, Lee JC, et al. (2000) Geologic knowledge and seismic risk mitigation: Insight from the Chi-Chi earthquake – Taiwan. In: Lo C-H and Liao W-I (eds.) *Proceedings of International Workshop on Annual Commemoration of Chi-Chi Earthquake, Science Aspect*, pp. 13–24.

Arias A (1970) A measure of earthquake intensity. In: Hansen RJ (ed.) *Seismic Design for Nuclear Power Plant*, pp. 438–483. Cambridge, MA: Massachusetts Institute of Technology.

Bakun W and Lindh A (1985) The Parkfield earthquake prediction experiment. *Science* 229: 619–624.

Bakun WH, Aagaard B, Dost B, et al. (2005) Implications for prediction and hazard assessment from the 2004 Parkfield earthquake. *Nature* 437: 969–974.

Bardet JP (2003) Advances in analysis of soil liquefaction during earthquakes. In: Lee WHK, Kanamori H, Jennings PC, and Kisslinger C (eds.) *International Handbook of Earthquake and Engineering Seismology, Part B*, pp. 1175–1201. San Diego, CA: Academic Press.

Beresnev IA (2006) Theory of vibratory mobilization of nonwetting fluids entrapped in pore constrictions. *Geophysics* 71: N47–N56.

Beresnev IA, Gaul W, and Vigil RD (2011) Direct pore-level observation of permeability increase in two-phase flow by shaking. *Geophysical Research Letters* 38: L20302.

Berryman JG and Wang HF (2000) Elastic wave propagation and attenuation in a double-porosity dual-permeability medium. *International Journal of Rock Mechanics and Mining Sciences* 37: 67–78.

Biagi PF, Ermini A, Kingsley SP, Khatkevich YM, and Gordeev EI (2001) Difficulties with interpreting changes in groundwater gas content as earthquake precursors in Kamchatka, Russia. *Journal of Seismology* 5: 487–497.

Biot MA (1941) General theory of three-dimensional consolidation. *Journal of Applied Physics* 12: 155–164.

Biot MA (1956a) Theory of propagation of elastic waves in fluid-saturated porous solid. I. Low-frequency range. *Journal of the Acoustical Society of America* 28: 168–178.

Biot MA (1956b) Theory of propagation of elastic waves in fluid-saturated porous solid. I. Higher frequency range. *Journal of the Acoustical Society of America* 28: 179–191.

Biot MA (1962) Mechanics of deformation and acoustic propagation in porous media. *Journal of Applied Physics* 33: 1482–1498.

Bird P (1978) Stress and temperature in subduction shear zones: Tonga and Mariana. *Geophysical Journal of the Royal Astronomical Society* 55: 411–434.

Blanchard FB and Byerly P (1935) A study of a well gauge as a seismograph. *Bulletin of the Seismological Society of America* 25: 313–321.

Bonini M (2009) Mud volcano eruptions and earthquakes in the Northern Apennines and Sicily, Italy. *Tectonophysics* 474: 723–735.

Bosl WJ and Nur A (2002) Aftershocks and pore fluid diffusion following the 1992 Landers earthquake. *Journal of Geophysical Research* 107: 2366.

Bower DR and Heaton KC (1978) Response of an aquifer near Ottawa to tidal forcing and the Alaskan earthquake of 1964. *Canadian Journal of Earth Sciences* 15: 331–340.

Brace WF, Paulding J, and Scholz C (1966) Dilatancy in the fracture of crystalline rocks. *Journal of Geophysical Research* 71: 3939–3953.

Briggs RO (1991) Effects of Loma Prieta earthquake on surface water in Waddell Valley. *Water Resources Bulletin* 27: 991–999.

Brodsky EE and Kanamori H (2001) Elastohydrodynamic lubrication of faults. *Journal of Geophysical Research* 106: 16357–16374.

Brodsky EE and Prejean SG (2005) New constraints on mechanisms of remotely triggered seismicity at Long Valley Caldera. *Journal of Geophysical Research* 110. http://dx.doi.org/10.1029/2004JB003211.

Brodsky EE, Roeloffs E, Woodcock D, Gall I, and Manga M (2003) A mechanism for sustained ground water pressure changes induced by distant earthquakes. *Journal of Geophysical Research* 108. http://dx.doi.org/10.1029/2002JB002321.

Brodsky EE, Mori J, and Fulton PM (2010) Drilling into faults quickly after earthquakes. *Eos, Transactions American Geophysical Union* 91: 23744. http://dx.doi.org/10.1029/2010EO270001.

Brown KM, Tryon MD, DeShon HR, Schwartz S, and Dorman LM (2005) Correlated transient fluid pulsing and seismic tremor in the Costa Rica subduction zone. *Earth and Planetary Science Letters* 238: 189–203.

Camacho A, Lee JKW, Hensen BJ, and Braun J (2005) Short-lived orogenic cycles and the eclogitization of cold crust by spasmodic hot fluids. *Nature* 435: 1191–1196.

Cappa F, Rutqvist J, and Yamamoto K (2009) Modeling crustal deformation and rupture processes related to upwelling of deep CO_2-rich fluids during the 1965–1967 Matsushiro earthquake swarm in Japan. *Journal of Geophysical Research* 114: B10304. http://dx.doi.org/10.1029/2009JB006398.

Carson B and Screaton EJ (1998) Fluid flow in accretionary prisms: Evidence for focused, time-variable discharge. *Reviews of Geophysics* 36: 329–352.

Chen CH, et al. (2010) Identification of earthquake signals from groundwater level records using the HHT method. *Geophysical Journal International* 180: 1231–1241.

Chia YP, et al. (2008) Implications of coseismic groundwater level changes observed at multiple-well monitoring stations. *Geophysical Journal International* 172: 293–301.

Charmoille A, Fabbri O, Mudry J, Guglielmi Y, and Bertrand C (2005) Post-seismic permeability change in a shallow fractured aquifer following a M-L 5.1 earthquake (Fourbanne karst aquifer, Jura outermost thrust unit, eastern France). *Geophysical Research Letters* 32: L18406.

Chester FM, Evens JP, and Biegel RL (1993) Internal structure and weakening mechanisms of the San Andreas fault. *Journal of Geophysical Research* 98: 771–786.

Chester FM, Chester JS, Kirschner DL, Schulz SE, and Evans JP (2004) Structure of large-displacement strike-slip fault zones in the brittle continental crust. In: Karner GD, Taylor B, Driscoll NW, and Kohlstedt DL (eds.) *Rheology and Deformation in the Lithosphere at Continental Margins*, pp. 223–260. New York: Columbia University Press.

Chia YP, Wang YS, Wu HP, Chiu JJ, and Liu CW (2001) Changes of groundwater level due to the 1999 Chi-Chi earthquake in the Choshui River fan in Taiwan. *Bulletin of the Seismological Society of America* 91: 1062–1068.

Christiansen LB, Hurwitz S, and Ingebritsen SE (2007) Annual modulation of seismicity along the San Andreas Fault near Parkfield. *Geophysical Research Letters* 34: L04306. http://dx.doi.org/10.1029/2006GL028634.

Christiansen LB, Hurwitz S, Saar MO, Ingebritsen SE, and Hsieh PA (2005) Seasonal seismicity at western United States volcanic centers. *Earth and Planetary Science Letters* 240: 307–321.

Chu C-L, Wang C-Y, and Lin W (1981) Permeability and frictional properties of San Andreas fault gouges. *Geophysical Research Letters* 8: 56568. http://dx.doi.org/10.1029/GL008i006p00565.

Claesson L, Skelton A, Graham C, et al. (2004) Hydrogeochemical changes before and after a major earthquake. *Geology* 32L: 641–644.

Claesson L, Skelton A, Graham C, and Morth C-M (2007) The timescale and mechanisms of fault sealing and water–rock interaction after an earthquake. *Geofluids* 7: 427–440.

Cochran ES, Vidale JE, and Tanaka S (2004) Earth tides can trigger shallow thrust fault earthquakes. *Science* 306: 1164–1166.

Cooper HH, Bredehoeft JD, Papadopulos IS, and Bennett RR (1965) The response of well-aquifer systems to seismic waves. *Journal of Geophysical Research* 70: 3915–3926.

Costain JK, Bollinger GA, and Speer JA (1987) Hydroseismicity: A hypothesis for the role of water in the generation of intraplate seismicity. *Seismological Research Letters* 58: 41–64.

Cox SC, Rutter HJ, Sims A, et al. (2012) Hydrological effects of the Darfield (Canterbury) Mw 7.1 earthquake, 4 September 2010, New Zealand. *New Zealand Journal of Geology and Geophysics* 55: 231–247.
do Nascimento AF, Lunn RJ, and Cowie PA (2005) Numerical modelling of pore-pressure diffusion in a reservoir-induced seismicity site in northeast Brazil. *Geophysical Journal International* 160: 249–262.
Doan ML and Cornet FH (2007) Small pressure drop triggered near a fault by small teleseismic waves. *Earth and Planetary Science Letters* 258: 207–218.
Dobry R (1989) Some basic aspects of soil liquefaction during earthquakes. *Annals of the New York Academy of Sciences* 558: 172–182.
Dvorkin J, Nolen-Hoeksema R, and Nur A (1994) The squirt-flow mechanism: Macroscopic description. *Geophysics* 59: 428–438.
Esposito E, Pece R, Porfido S, and Tranfaglia G (2009) Ground effects and hydrological changes in Southern Apennines (Italy) in response to the 23 July 1930 earthquake (Ms = 6.7). *Natural Hazards and Earth System Sciences* 9: 539–550.
Elkhoury JE, Brodsky EE, and Agnew DC (2006) Seismic waves increase permeability. *Nature* 441: 1135–1138.
Forster CB, Evens JP, Tanaka H, Jeffreys R, and Noraha T (2003) Hydrologic properties and structure of the Mozumi Fault, central Japan. *Journal of Geophysical Research* 30. http://dx.doi.org/10.1029/2002GL014904.
Galli P (2000) New empirical relationships between magnitude and distance for liquefaction. *Tectonophysics* 324: 169–187.
Gavrilenko P, MeliKadze G, Chelidze T, Gibert D, and Kumsiashvili G (2000) Permanent water level drop associated with the Spitak earthquake: Observations at Lisi borehole (Republic of Georgia) and modelling. *Geophysical Journal International* 143: 83–98.
Ge S, Liu M, Lu N, Godt LW, and Luo G (2009) Did the Zipingpu Reservoir trigger the 2008 Wenchuan earthquake? *Geophysical Research Letters* 36: L20315. http://dx.doi.org/10.1029/2009GL040349.
Geballe ZM, Wang CY, and Manga M (2011) A permeability-change model for water level changes triggered by teleseismic waves. *Geofluids* 11: 302–308.
Gomberg J and Wolf L (1999) A possible cause for an improbable earthquake: The 1997 Mw 4.9 southern Alabama earthquake and hydrocarbon recovery. *Geology* 27: 367–370.
Gorman P (2011) Scientists poised to drill into fault: Science/Stuff.co.nz (http://www.stuff.co.nz/science/4511807/Scientists-poised-to-drillinto-fault).
Gorokhovich Y (2005) Abandonment of Minoan palaces on Crete in relation to the earthquake induced changes in groundwater supply. *Journal of Archaeological Science* 32: 217–222.
Gudmundsson A (2000) Active fault zones and groundwater flow. *Geophysical Research Letters* 27: 2993–2996.
Guilhem A, Peng Z, and Nadeau RM (2010) High-frequency identification of non-volcanic tremor triggered by regional earthquakes. *Geophysical Research Letters* 37: L16309. http://dx.doi.org/10.1029/2010GL044660.
Gupta HK (1992) *Reservoir-Induced Earthquakes*, 264 pp. New York: Elsevier.
Hainzl S (2004) Seismicity patterns of earthquake swarms due to fluid intrusion and stress triggering. *Geophysical Journal International* 159: 1090–1096.
Hainzl S, Kraft T, Wassermann J, Igel I, and Schmedes E (2006) Evidence for rainfall-triggered earthquake activity. *Geophysical Research Letters* 33: L19303. http://dx.doi.org/10.1029/2006GL027642.
Hamiel Y, Lyakhovsky V, and Agnon A (2005) Rock dilation, nonlinear deformation, and pore pressure change under shear. *Earth and Planetary Science Letters* 237: 577–589.
Hancox GT, Cox SC, Turnbull IM, and Crozier MJ (2003) Landslides and other ground damage caused by the Mw 7.2 Miordland earthquake of 22 August 2003. Institute of Geological and Nuclear Sciences, Science Report 2003/30, Lower Hutt.
Harris RA (1998) Introduction to the special section: Stress triggers, stress shadows, and the implication for seismic hazard. *Journal of Geophysical Research* 103: 24247–24258.
Hartman J and Levy JK (2005) Hydrogeological and gasgeochemical earthquake precursors – A review for application. *Natural Hazards* 34: 279–304.
Heki K (2003) Snow load and seasonal variation of earthquake occurrence in Japan. *Earth and Planetary Science Letters* 207: 159–164.
Holzer TL and Youd TL (2007) Liquefaction, ground oscillation, and soil deformation at the Wildlife Array, California. *Bulletin of the Seismological Society of America* 97: 961–976.
Hsu SK, Chin CV, Cheng LH, and Lin WS (1999) The change of groundwater level of earthquake 921 in central Taiwan. In: *Proceedings of Conference on Sustainable Development and Geosciences Prospect: National Taiwan University, Department of Geology,* pp. 7–35 (in Chinese).
Husen S and Kissling E (2001) Postseismic fluid flow after the large subduction earthquake of Antofagasta, Chile. *Geology* 29: 847–850.
Husen S, Taylor R, Smith RB, and Healser H (2004) Changes in geyser eruption behavior and remotely triggered seismicity in Yellowstone National Park produced by the 2002 M 7.9 Denali fault earthquake, Alaska. *Geology* 32: 537–540.
Hutchinson RA (1985) Hydrothermal changes in the upper geyser basin, Yellowstone National Park, after the 1983 Borah Peak, Idaho, earthquake. USGS Open File Report 85-290, 612–624.
Hutchinson RA, Westphal JA, and Kieffer SW (1997) In situ observations of Old Faithful Geyser. *Geology* 25: 875–878.
Igarashi G and Wakita H (1991) Tidal responses and earthquake-related changes in the water level of deep wells. *Journal of Geophysical Research* 96: 4269–4278.
Igarashi G and Wakita H (1995) Geochemical and hydrological observations for earthquake prediction in Japan. *Journal of Physics of the Earth* 43: 585–598.
Igarashi G, Saeki S, Takahata N, et al. (1995) Ground-water radon anomaly before the Kobe earthquake in Japan. *Science* 269: 60–61.
Ingebritsen SE and Rojstaczer SA (1993) Controls on geyser periodicity. *Science* 262: 889–892.
Ingebritsen SE, Sanford WE, and Neuzil CE (2006) *Groundwater in Geologic Processes*, 2nd edn., 562 pp. New York: Cambridge University Press.
Ishihara K (1996) *Soil Behavior in Earthquake Geotechnics*, 350 pp. Oxford: Clarendon Press.
Itaba S and Koizumi N (2007) Earthquake-related changes in groundwater levels at the Dogo hot spring, Japan. *Pure and Applied Geophysics* 164: 2397–2410.
Jennings PC (2003) An introduction to the earthquake response of structures. In: Lee WHK, Kanamori H, Jennings PC, and Kisslinger C (eds.) *International Handbook of Earthquake and Engineering Seismology, Part B*, pp. 1097–1126. London: Academic Press.
Jonsson S, Segall P, Pedersen R, and Bjornsson G (2003) Post-earthquake ground movements correlated to pore-pressure transients. *Nature* 424: 179–183.
King C-Y, Koizumi N, and Kitagawa Y (1995) Hydrogeochemical anomalies and the 1995 Kobe earthquake. *Science* 269: 38–39.
King C-Y, Azuma S, Igarashi G, Ohno M, Saito H, and Wakita H (1999) Earthquake-related water-level changes at 16 closely clustered wells in Tono, central Japan. *Journal of Geophysical Research* 104: 13073–13082.
King C-Y, Azuma S, Ohno M, et al. (2000) In search of earthquake precursors in the water-level data of 16 closely clustered wells at Tono, Japan. *Geophysical Journal International* 143: 469–477.
King CY, Zhang W, and Zhang ZC (2006) Earthquake-induced groundwater and gas changes. *Pure and Applied Geophysics* 163: 633–645.
Kingsley SP, Biagi PF, Piccolo R, et al. (2001) Hydrogeochemical precursors of strong earthquakes: A realistic possibility in Kamchatka. *Physics and Chemistry of the Earth, Part C* 26: 769–774.
Kitagawa Y, Fujimori K, and Koizumi N (2007) Temporal change in permeability of the Nojima Fault zone by repeated water injection experiments. *Tectonophysics* 443: 18392.
Kitagawa Y, Koizumi N, Takahashi M, Matsumoto N, and Sato T (2006) Changes in groundwater levels or pressures associated with the 2004 earthquake off the west coast of northern Sumatra (M 9.0). *Earth, Planets and Space* 58: 173–179.
Koizumi N, Lai W-C, Kitagawa Y, and Matsumoto N (2004) Comment on "Coseismic hydrological changes associated with dislocation of the September 21, 1999 Chichi earthquake, Taiwan" In: Lee M, et al. (eds.) *Geophysical Research Letters* 31. http://dx.doi.org/10.1029/2004GL019897.
Kono Y and Yanagidani T (2006) Broadband hydroseismograms observed by closed borehole wells in the Makioka mine, central Japan: Response of pore pressure to seismic waves from 0.05 to 2 Hz. *Journal of Geophysical Research* 111: B03410.
Kopf AJ (2002) Significance of mud volcanism. *Reviews of Geophysics* 40. http://dx.doi.org/10:1029/2000RG000093.
Kuribayashi E and Tatsuoka F (1975) Brief review of liquefaction during earthquakes in Japan. *Soils and Foundations* 15: 81–92.
Lachenbruch A (1980) Frictional heating, fluid pressure, and the resistance to fault motion. *Journal of Geophysical Research* 85: 6097–6112.
Lachenbruch AH and Sass JH (1977) Heat flow in the United States and the thermal regime of the crust. In: Heacock JH (ed.) *The Earth's Crust. Geophysical Monograph Series*, vol. 20, pp. 625–675. Washington, DC: American Geophysical Union.
Lawson AC (1908) The California Earthquake of April 18, 1906. *Report of the State Earthquake Investigation Commission*, vol. 1. Washington, DC.: Carnegie Institution of Washington.
Lay T and Wallace TC (1995) *Modern Global Seismology*. Academic Press.
Lee CS and Wu CL (1996) Pumping tests of the Choshuichi alluvial fan. In: *Conference on Groundwater and Hydrology of the Choshui River Alluvial Fan, Taiwan, Water Resources Bureau,* pp. 165–179 (in Chinese).
Lee CT, Kelson KI, and Kang KH (2000) Hanging wall deformation and its effect on buildings and structures as learned from the Chelungpu faulting in the 1999

Chi-Chi Taiwan earthquake. In: Lo C-H and Liao W-I (eds.) *Proceedings of International Workshop Annual Commemoration of Chi-Chi Earthquake, Science Aspect*, pp. 93–104.

Lee J-C, Chu H-T, Angelier J, et al. (2002a) Geometry and structure of northern rupture surface ruptures of the 1999 Mw = 7.6 Chi-Chi Taiwan earthquake: Influence from inherited fold belt structures. *Journal of Structural Geology* 24: 173–192.

Lee M, Liu T-K, Ma K-F, and Chang Y-M (2002) Coseismic hydrological changes associated with dislocation of the September 21, 1999 Chi-Chi earthquake, Taiwan. *Geophysical Research Letters* 29. http://dx.doi.org/10.1029/2002GL015116.

Li Y-G and Malin PE (2008) San Andreas Fault damage at SAFOD viewed with fault-guided waves. *Geophysical Research Letters* 35: L08304. http://dx.doi.org/10.1029/2007GL032924.

Li H and Zhu L (2007) High-resolution structures of the Landers fault zone inferred from aftershock waveform data. *Geophysical Journal International* 171: 1295–1307.

Li Y-G, Vidale JE, Aki K, and Xu F (2000) Depth-dependent structure of the Landers Fault zone from trapped waves generated by aftershocks. *Journal of Geophysical Research* 105: 6237–6254.

Lin WY (2000) Unstable groundwater supply after the big earthquake. *United Daily Newspaper*, Taiwan (in Chinese).

Linde AT and Sacks IS (2002) Slow earthquakes and great earthquakes along the Nankai trough. *Earth and Planetary Science Letters* 203: 265–275.

Linde AT, Sacks IS, Johnston MJS, Hill DP, and Bilham RG (1994) Increased pressure from rising bubbles as a mechanism for remotely triggered seismicity. *Nature* 371: 408–410.

Liu E, Vlatsos S, Li S-Y, Main IG, and Schoenberg M (2004) Modeling seismic wave propagation during fluid injection in a fractured network: Effects of pore fluid pressure on time-lapse seismic signatures. *Leading Edge* 23: 778–783.

Liu LB, Roeloffs E, and Zheng XY (1989) Seismically induced water level oscillations in the Wali well, Beijing, China. *Journal of Geophysical Research* 94: 9453–9462.

Lo T, Coyner KB, and Toksoz N (1986) Experimental determination of elastic anisotropy of Berea sandstone, Chicopee shale, and Chelmsford granite. *Geophysics* 51: 16471. http://dx.doi.org/10.1190/1.1442029.

Lockner DA and Beeler NM (2002) Rock failure and earthquakes. In: Lee WHK, Kanamori H, Jennings PC, and Kisslinger C (eds.) *International Handbook of Earthquake and Engineering Seismology*, pp. 505–537. San Diego, CA: Academic Press.

Ma KF, Lee CT, and Tsai YB (1999) The Chi-Chi, Taiwan earthquake: Large surface displacement on an inland thrust fault. *Eos, Transactions American Geophysical Union* 80: 605–611.

Maeda K and Yoshida A (1990) A probabilistic estimation of earthquake occurrence on the basis of the appearance times of multiple precursory phenomena. *Journal of Physics of the Earth* 38: 431–444.

Maltman AJ and Bolton A (2003) How sediments become mobilized. In: Van Rensberger P, et al. (eds.) *Subsurface Sediment Mobilization*, *Geological Society of London Special Publications*, vol. 216, pp. 9–20.

Manga M (2001) Origin of postseismic streamflow changes inferred from baseflow recession and magnitude-distance relations. *Geophysical Research Letters* 28: 2133–2136.

Manga M and Bonini M (2012) Large historical eruptions at subaerial mud volcanoes, Italy. *Natural Hazards and Earth System Sciences* 12: 3377–3386.

Manga M and Brodsky EE (2006) Seismic triggering of eruptions in the far field: Volcanoes and geysers. *Annual Review of Earth and Planetary Sciences* 34: 263–291.

Manga M, Brodsky EE, and Boone M (2003) Response of streamflow to multiple earthquakes and implications for the origin of postseismic discharge changes. *Geophysical Research Letters* 30. http://dx.doi.org/10.1029/2002GL016618.

Manga M, Brumm M, and Rudolph ML (2009) Earthquake triggering of mud volcanoes. *Marine and Petroleum Geology* 26: 1785–1798.

Manga M and Rowland JC (2009) Response of Alum Rock springs to the October 30, 2007 earthquake and implications for the origin of increased discharge after earthquakes. *Geofluids* 9: 237–250.

Manga M, Beresnev I, Brodsky EE, et al. (2012) Changes in permeability by transient stresses: Field observations, experiments and mechanisms. *Reviews of Geophysics* 50: RG2004. http://dx.doi.org/10.1029/2011RG000382.

Manning CE and Ingebritsen SE (1999) Permeability of the continental crust: Constraints from heat flow models and metamorphic systems. *Reviews of Geophysics* 37: 127–150.

Marone C, Carpenter BM, Rathbun AP, and Saffer DM (2010) Physical properties and mechanical behavior of the active San Andres Fault Zone: Insight from laboratory studies. *EOS Transactions American Geophysical Union*, abstract T52B04.

Mase CW and Smith L (1987) Effect of frictional heating on the thermal, hydrological, and mechanical response of a fault. *Journal of Geophysical Research* 92: 6249–6272.

Masson Y and Pride SR (2006) Poroelastic finite-difference modeling of seismic attenuation and dispersion due to mesoscopic-scale heterogeneity. *Journal of Geophysical Research* 112: B03204.

Matsumoto N (1992) Regression analysis for anomalous changes of ground water level due to earthquakes. *Geophysical Research Letters* 19: 1192–1196.

Matsumoto N and Roeloffs EA (2003) Hydrological response to earthquakes in the Haibara well, central Japan—II. Possible mechanism inferred from time-varying hydraulic properties. *Geophysical Journal International* 155: 899–913.

Mavko G, Mukerji T, and Dvorkin J (1998) *The Rock Physics Handbook: Tools for Seismic Data Analysis in Porous Media*. New York: Cambridge University Press.

Mazzini A, Svensen H, Akhmanov GG, et al. (2007) Triggering and dynamic evolution of the LUSI mud volcano, Indonesia. *Earth and Planetary Science Letters* 261: 375–388.

Mazzini A, Nermoen A, Krotkiewski M, Podladchikov Y, Planke S, and Svensen H (2009) Strike-slip faulting as a trigger mechanism for overpressure release through piercement structures. Implications for the Lusi mud volcano, Indonesia. *Marine and Petroleum Geology* 26: 1751–1765.

McEvilly T and Johnson L (1974) Stability of P and S velocities from central California quarry blasts. *Bulletin of the Seismological Society of America* 64: 343–353.

Mellors R, Kilb D, Aliyev A, Gasanov A, and Yetirmishli G (2007) Correlations between earthquakes and large mud volcano eruptions. *Journal of Geophysical Research* 112: B04304.

Micklethwaite S and Cox SF (2004) Fault-segment rupture, aftershock-zone fluid flow, and mineralization. *Geology* 32: 813–816.

Milkov AV (2000) Worldwide distribution of submarine mud volcanoes and associated gas hydrates. *Marine Geology* 167: 29–42.

Miller SA, Collettini C, Chiaraluce L, Cocco M, Barchi M, and Kaus BJP (2004) Aftershocks driven by a high-pressure CO_2 source at depth. *Nature* 427: 724–727.

Mizuno T (2003) The subsurface observation of fault-zone trapped waves: Applications to investigations of the deep structure of active faults. *Bulletin of the Earthquake Research Institute* 78: 91–106.

Mizuno T and Kuwahara Y (2008) Spatial variations in fault-zone structure along the Nojima fault, central Japan, as inferred from borehole observations of fault-zone trapped waves. *Bulletin of the Seismological Society of America* 98: 558–570.

Mogi K, Mochizuki H, and Kurokawa Y (1989) Temperature changes in an artesian spring at Usami in the Izu Peninsula (Japan) and their relation to earthquakes. *Tectonophysics* 159: 95–108.

Mohr CH, Montgomery DR, Huber A, Bronstert A, and Iroum A (2012) Streamflow response in small upland catchments in the Chilean coastal range to the MW 8.8 Maule earthquake on 27 February 2010. *Journal of Geophysical Research* 117: F02032. http://dx.doi.org/10.1029/2011JF002138.

Montgomery DR and Manga M (2003) Streamflow and water well responses to earthquakes. *Science* 300: 2047–2049.

Montgomery DR, Greenberg HM, and Smith DT (2003) Streamflow response to the Nisqually earthquake. *Earth and Planetary Science Letters* 209: 19–28.

Morrow C, Shi LQ, and Byerlee J (1984) Permeability of fault gouge under confining pressure and shear stress. *Journal of Geophysical Research* 89: 3193–3200. http://dx.doi.org/10.1029/JB089iB05p03193.

Mount VS and Suppe J (1987) State of stress near the San Andreas fault—Implications doe wrench tectonics. *Geology* 15: 1143–1146.

Muir-Wood R and King GCP (1993) Hydrological signatures of earthquake strain. *Journal of Geophysical Research* 98: 22035–22068.

National Research Council (1985) *Liquefaction of Soils During Earthquakes*, 240 pp. Washington, DC: National Academy Press.

Neuzil CE (2003) Hydromechanical coupling in geologic processes. *Hydrogeology Journal* 11: 41–83.

Nur A (1974) Matsushiro, Japan, earthquake swarm: Confirmation of the dilatancy-fluid diffusion model. *Geology* 2: 217–221.

Nur A and Booker JR (1972) Aftershocks caused by pore fluid flow? *Science* 175: 885–887.

Ohtani T, Tanaka H, Fujimoto K, Higuchi T, Tomida N, and Ito H (2001) Internal structure of the Nojima fault zone from Hirabayashi GSJ drill core. *The Island Arc* 10: 392–400.

Panahi BM (2005) Mud volcanism, geodynamics and seismicity of Azerbaijan and the Caspian Sea region. In: Martinelli G and Panahi B (eds.) *Mud Volcanoes, Geodynamics and Seismicity*, pp. 89–104. Dordrecht, The Netherlands: Springer.

Papadopoulos GA and Lefkopulos G (1993) Magnitude-distance relations for liquefaction in soil from earthquakes. *Bulletin of the Seismological Society of America* 83: 925–938.

Park SK (1997) Monitoring resistivity changes in Parkfield, California: 1988–1995. *Journal of Geophysical Research* 102: 24545–24559.

Park SK (2005) The agony and ecstasy of earthquake prediction (talk presented at U.S. Geological Survey, Menlo Park, CA).

Parotidis M and Shapiro SA (2004) A statistical model for the seismicity rate of fluid-injection-induced earthquakes. *Geophysical Research Letters* 31. http://dx.doi.org/10.1029/2004GL020421.

Plastino W, Povinec PP, De Luca G, et al. (2010) Uranium groundwater anomalies and L'Aquila earthquake, 6th April 2009 (Italy). *Journal of Environmental Radioactivity* 101: 45–50.

Pralle N, Külzer M, and Gudehus G (2003) Experimental evidence on the role of gas in sediment liquefaction and mud volcanism. In: Van Rensberger P, et al. (eds.) *Subsurface Sediment Mobilization*, *Geological Society of London Special Publications*, 216: 159–172.

Pride SR, Berryman JG, and Harris JM (2004) Seismic attenuation due to wave-induced flow. *Journal of Geophysical Research* 109: B01201. http://dx.doi.org/10.1029/2003JB002639.

Quigley MC, Bastin S, and Bdadley BA (2013) Recurrent liquefaction in Christchurch, New Zealand, during the Canterbury earthquake sequence. *Geology.* http://dx.doi.org/10.1130/G33944.1.

Quilty EG and Roeloffs EA (1997) Water level changes in response to the December 20, 1994 M4.7 earthquake near Parkfield, California. *Bulletin of the Seismological Society of America* 87: 1018–1040.

Raleigh CB, Healy JH, and Bredehoeft JD (1976) An experiment in earthquake control at Rangely, Colorado. *Science* 191: 1230–1237.

Rathbun AP, Song I, and Saffer D (2010) Permeability of the San Andreas fault zone at depth. *EOS Transactions American Geophysical Union*, abstract T41A2095.

Rice JR (1992) Fault stress state, pore pressure distributions, and the weakness of the san Andreas fault. In: Evans B and Wong T-F (eds.) *Fault Mechanics and Transport Properties of Rocks*, pp. 475–504. San Diego, CA: Academic Press.

Rice JR and Cleary MP (1976) Some basic stress diffusion solutions for fluid saturated elastic porous media with compressible constituents. *Reviews of Geophysics and Space Physics* 14: 227–241.

Rinehart JS (1972) Fluctuations in geyser activity caused by variations in earth tidal forces, barometric pressure, and tectonic stresses. *Journal of Geophysical Research* 77: 342–350.

Roeloffs EA (1996) Poroelastic techniques in the study of earthquake-related hydrologic phenomena. *Advances in Geophysics* 37: 135–195.

Roeloffs EA (1998) Persistent water level changes in a well near Parkfield, California, due to local and distant earthquakes. *Journal of Geophysical Research* 103: 869–889.

Roeloffs EA (2006) Evidence for aseismic deformation-rate changes prior to earthquakes. *Annual Review of Earth and Planetary Sciences* 34: 591–627.

Roeloffs EA and Quilty E (1997) Water level and strain changes preceding and following the August 4, 1985 Kettleman Hills, California, earthquake. *Pure and Applied Geophysics* 149: 21–60.

Roeloffs EA, Sneed M, Galloway DL, et al. (2003) Water-level changes induced by local and distant earthquakes at Long Valley caldera, California. *Journal of Volcanology and Geothermal Research* 127: 269–303.

Rojstaczer S, Ingebritsen SE, and Hayba DO (2005) Permeability of continental crust influenced by internal and external loading. *Geofluids* 8: 128–139.

Rojstaczer SA, Ingebritsen SE, and Hayba DO (2008) Permeability of continental crust influenced by internal and external forcing. *Geofluids* 8: 128–139. http://dx.doi.org/10.1111/j.1468-8123.2008.00211.x.

Rojstaczer S and Wolf S (1992) Permeability changes associated with large earthquakes: An example from Loma Prieta, California. *Geology* 20: 211–214.

Rojstaczer S, Wolf S, and Michel R (1995) Permeability enhancement in the shallow crust as a cause of earthquake-induced hydrological changes. *Nature* 373: 237–239.

Rojstaczer S, Galloway DL, Ingebritsen SE, and Rubin DM (2003) Variability in geyser eruptive timing and its causes: Yellowstone National Park. *Geophysical Research Letters* 30: Art. No. 1953.

Rudnicki JW, Yin J, and Roeloffs EA (1993) Analysis of water level changes induced by fault creep at Parkfield, California. *Journal of Geophysical Research* 98: 8143–8152.

Rudolph ML and Manga M (2010) Mud volcano response to the April 4, 2010 El Mayor–Cucapah earthquake. *Journal of Geophysical Research* 115: B12211. http://dx.doi.org/10.1029/2010JB007737.

Rudolph ML and Manga M (2012) Frequency dependence of mud volcano response to earthquakes. *Geophysical Research Letters* 39: L14303. http://dx.doi.org/10.1029/2012GL052383.

Saar MO and Manga M (2003) Seismicity induced by seasonal groundwater recharge at Mt. Hood, Oregon. *Earth and Planetary Science Letters* 214: 605–618.

Sato H (1978) Precursory land tilt prior to the Tonankai earthquake of 1944. In some precursors prior to recent great earthquakes along the Nankai trough. *Journal of Physics of the Earth* 25: s115–s121.

Sato H (1982) On the changes in the sea level at Tosashimizu before the Nankaido earthquake of 1946. *Journal of the Seismological Society of Japan* 35: 623–626.

Sato T, Sakai R, Furuya K, and Kodama T (2000) Coseismic spring flow changes associated with the 1995 Kobe earthquake. *Geophysical Research Letters* 27: 1219–1222.

Sato T, Matsumoto N, Kitagawa Y, et al. (2004) Changes in water level associated with the 2003 Tokachi-oki earthquake. *Earth, Planets and Space* 56: 395–400.

Savarensky EF (1968) On the prediction of earthquakes. *Tectonophysics* 6: 17–27.

Scholtz CH, Sykes LR, and Aggarwal YP (1973) Earthquake prediction—Physical basis. *Science* 181: 803–810.

Schuster RL and Murphy W (1996) Structural damage, ground failure, and hydrological effects of the magnitude (Mw) 5.9 Draney Peak, Idaho, earthquake of February 3, 1994. *Seismological Research Letters* 67: 20–29.

Seed H and Lee KL (1966) Liquefaction of saturated sands during cyclic loading. *Journal of the Soil Mechanics and Foundations Division* 92: 105–134.

Segall P (1989) Earthquakes triggered by fluid extraction. *Geology* 17: 942–946.

Segall P, Grasso JR, and Mossop A (1994) Poroelastic stressing and induced seismicity near the Lacq gas field, southwestern France. *Journal of Geophysical Research* 99: 15423–15438.

Semenov AM (1969) Variations in the travel-time of transverse and longitudinal waves before violent earthquakes. *Physics of the Solid Earth* 3: 245–248.

Shapiro SA, et al. (2006) Fluid induced seismicity guided by a continental fault: Injection experiment of 2004/2005 at the German Deep Drilling Site (KTB). *Geophysical Research Letters* 33: L01309. http://dx.doi.org/10.1029/2005GL024659.

Shi Z, Wang G, Wang C-Y, Manga M, and Liu C (2014) Comparison of hydrological responses to the Wenchuan and Lushan earthquakes. *Earth and Planetary Science Letters* 391: 193–200.

Sibson RH (1973) Interactions between temperature and pore-fluid pressure during earthquake faulting and a mechanism for partial or total stress relief. *Nature* 243: 66–68.

Sibson RH and Rowland JV (2003) Stress, fluid pressure and structural permeability in seismogenic crust, North Island, New Zealand. *Geophysical Journal International* 154: 584–594.

Sil S and Freymueller JT (2006) Well water level changes in Fairbanks, Alaska, due to the great Sumatra–Andaman earthquake. *Earth, Planets and Space* 58: 181–184.

Silver PG and Valette-Silver NJ (1992) Detection of hydrothermal precursors to large northern California earthquakes. *Science* 257: 1363–1368.

Simpson DW, Leith WS, and Scholz CH (1988) Two types of reservoir-induced seismicity. *Bulletin of the Seismological Society of America* 78: 2025–2050.

Sleep NH and Blanpied ML (1992) Creep, compaction and the weak rheology of major faults. *Nature* 359: 687–692.

Sorey ML and Clark MD (1981) Changes in the discharge characteristics of thermal springs and fumaroles in the Long Valley caldera, California, resulting from earthquakes on May 25–27, 1980. US Geological Survey Open-File Report 81-203.

Stein RS (1999) The role of stress transfer in earthquake occurrence. *Nature* 402: 605–609.

Steinberg GS, Merzhanov AG, and Steinberg AS (1982a) Geyser process: Its theory, modeling and field experiment. Part 1. Theory of the geyser process. *Modern Geology* 8: 67–70.

Steinberg GS, Merzhanov AG, and Steinberg AS (1982b) Geyser process: Its theory, modeling and field experiment. Part 3. On metastability of water in geysers. *Modern Geology* 8: 75–78.

Steinberg GS, Merzhanov AG, and Steinberg AS (1982c) Geyser process: Its theory, modeling and field experiment. Part 4. On seismic influence on geyser regime. *Modern Geology* 8: 79–86.

Stidham C, Antolik M, Dreger D, Larsen S, and Romanowicz B (1999) Three-dimensional structure influences on the strong-motion wavefield of the 1989 Loma Prieta earthquake. *Bulletin of the Seismological Society of America* 89: 1184–1202.

Su T-C, Chiang K-W, Lin S-J, Wang F-G, and Duann S-W (2000) Field reconnaissance and preliminary assessment of liquefaction in Yuan-Lin area. *Sino-Geotechnics* 77: 29–38.

Tadokoro K, Ando M, and Nishigami K (2000) Induced earthquakes accompanying the water injection experiment at the Nojima fault zone, Japan: Seismicity and its migration. *Journal of Geophysical Research* 105: 6089–6104.

Talwani P and Acree S (1985) Pore pressure diffusion and the mechanism of reservoir-induced seismicity. *Pure and Applied Geophysics* 122: 947–965.

Terzaghi K, Peck RB, and Mesri G (1996) *Soil Mechanics in Engineering Practice*, 3rd edn. New York: Wiley.

Tsunogai U, Maegawa K, Sato S, et al. (2012) Coseismic massive methane release from a submarine mud volcano. *Earth and Planetary Science Letters* 341–344: 79–85.

Thurber C, Roecker S, Ellsworth W, Chen Y, Lutter W, and Sessions R (1997) Two-dimensional seismic image of the San Andreas fault in the Northern Gabilan range, Central California: Evidence for fluids in the fault zone. *Geophysical Research Letters* 24: 1591–1594.

Tokunaga T (1999) Modeling of earthquake-induced hydrological changes and possible permeability enhancement due to 17 January 1995 Kobe earthquake, Japan. *Journal of Hydrology* 223: 221–229.

Townend J (1997) Subducting a sponge; minimum estimates of the fluid budget of the Hikurangi Margin accretionary prism. *Geological Society of New Zealand Newsletter* 112: 14–16.

Trique M, Richon P, Perrier F, Avouac JP, and Sabroux JC (1999) Radon emanation and electric potential variations associated with transient deformation near reservoir lakes. *Nature* 399: 137–141.

Tsunogai U and Wakita H (1995) Precursory chemical changes in ground water: Kobe earthquake, Japan. *Science* 269: 61–62.

Unsworth M, Bedrosian P, Eisel M, Egbert G, and Siripunvaraporn W (2000) Along strike variations in the electrical structure of the San Andreas Fault at Parkfield, California. *Geophysical Research Letters* 27: 3021–3024.

Wakita H (1975) Water wells as possible indicators of tectonic strain. *Science* 189: 553–555.

Wakita H (1996) Geochemical challenge to earthquake prediction. *Proceedings of the National Academy of Sciences* 93: 3781–3786.

Wakita H, Nakamura Y, and Sano Y (1988) Short-term and intermediate-term geochemical precursors. *Pure and Applied Geophysics* 126: 78–89.

Wang C-H, Wang CY, Kuo C-H, and Chen W-F (2005a) Some isotopic and hydrological changes associated with the 1999 Chi-Chi earthquake, Taiwan. *Island Arc* 14: 37–54.

Wang C-Y (1984) On the constitution of the San Andreas fault. *Journal of Geophysical Research* 89: 5858–5866.

Wang C-Y (2007) Liquefaction beyond the near field. *Seismological Research Letters* 78: 512–517.

Wang C-Y (2011) High pore pressure, or its absence, in the San Andreas fault. *Geology* 39: 1047–1050. http://dx.doi.org/10.1130/G32294.1.

Wang C-Y and Chia Y (2008) Mechanisms of water level changes during earthquakes: Near field versus intermediate field. *Geophysical Research Letters* 35: L12402. http://dx.doi.org/10.1029/2008GL034227.

Wang C-Y and Manga M (2010a) *Earthquakes and Water*. *Lecture Notes in Earth Sciences*, vol. 114. Springer.

Wang C-Y and Manga M (2010b) Hydrologic responses to earthquakes—A general metric. *Geofluids* 10: 206–216.

Wang C-Y, Wang C-H, and Kuo C-Y (2004a) Temporal change in groundwater level following the 1999 (Mw = 7.5) Chi-Chi earthquake, Taiwan. *Geofluids* 4: 210–220.

Wang C-Y, Wang C-H, and Manga M (2004b) Coseismic release of water from mountains: Evidence from the 1999 (Mw = 7.5) Chi-Chi, Taiwan, earthquake. *Geology* 32: 769–772.

Wang C-Y, Rui F, Yao Z, and Shi X (1986) Gravity anomaly and density structure of the San Andreas fault zone. *Pure and Applied Geophysics* 124: 127–140.

Wang C-Y, Cheng L-H, Chin C-V, and Yu S-B (2001) Coseismic hydrologic response of an alluvial fan to the 1999 Chi-Chi earthquake, Taiwan. *Geology* 29: 831–834.

Wang C-Y, Dreger DS, Wang C-H, Mayeri D, and Berryman JG (2003) Field relations among coseismic ground motion, water level change and liquefaction for the 1999 Chi-Chi (Mw = 7.5) earthquake, Taiwan. *Geophysical Research Letters* 17. http://dx.doi.org/10.1029/2003GL017601.

Wang C-Y, Dreger D, Manga D, and Wong A (2004c) Streamflow increase due to rupturing of hydrothermal reservoirs: Evidence from the 2003 San Simeon, California, earthquake. *Geophysical Research Letters* 31: L10502. http://dx.doi.org/10.1029/2004GL020124.

Wang C-Y, Manga M, Wang C-H, and Chen C-H (2012b) Transient change in groundwater temperature after earthquakes. *Geology* 40: 119–122.

Wang C-Y, Wang L-P, Manga M, Wang C-H, and Chen C-H (2013) Basin-scale transport of heat and fluid induced by earthquakes. *Geophysical Research Letters* 40. http://dx.doi.org/10.1002/grl.50738.

Wang C-Y, Wong A, Dreger D, and Manga M (2005c) Liquefaction limit during earthquakes and underground explosions—Implications on ground-motion attenuation. *Bulletin of the Seismological Society of America* 96: 355–363.

Wang R, Woith H, Milkereit C, and Zschau J (2004d) Modelling of hydrogeochemical anomalies induced by distant earthquakes. *Geophysical Journal International* 157: 717–726.

Wang HF (2000) *Theory of Linear Poroelasticity*, 287 pp. Princeton, NJ: Princeton University Press.

Wells DL and Coppersmith KJ (1994) New empirical relationships among magnitude, rupture length, rupture width, rupture area, and surface displacement. *Bulletin of the Seismological Society of America* 4: 964–1002.

Whitcomb JH, Garmany JD, and Anderson DL (1973) Earthquake prediction: Variation of seismic velocities before the San Fernando earthquake. *Science* 180: 632–635.

White DE (1967) Some principles of geyser activity, mainly from Steamboat Springs, Nevada. *American Journal of Science* 265: 644–684.

Wibberley C and Shimamoto T (2003) Internal structure and permeability of major strike-slip fault zones: The Median Tectonic Line in W. Mie Prefecture, S.W. Japan. *Journal of Structural Geology* 25: 59–78.

Wibberley CAJ and Shimamoto T (2005) Earthquake slip weakening and asperities explained by thermal pressurization. *Nature* 436: 689–692.

Wolf LW, Rowe CA, and Horner RB (1997) Periodic seismicity near Mt. Ogden on the Alaska-British Columbia border: A case for hydrologically-triggered earthquakes? *Bulletin of the Seismological Society of America* 87: 1473–1483.

Wong A and Wang CY (2007) Field relations between the spectral composition of ground motion and hydrological effects during the 1999 Chi-Chi (Taiwan) earthquake. *Journal of Geophysical Research* 112: B10305. http://dx.doi.org/10.1029/2006JB004516.

Wyss M and Booth DC (1997) The IASPEI procedure for the evaluation of earthquake precursors. *Geophysical Journal International* 131: 423–424.

Yan HR (2001) Water problems in the Yuenlin Mountains: Changes in groundwater after the Chi-Chi earthquake. *Chinese Daily,* Taiwan (in Chinese).

Youd TL (2003) Liquefaction mechanisms and induced ground failure. In: Lee WHK, Kanamori H, Jennings PC, and Kisslinger C (eds.) *International Handbook of Earthquake and Engineering Seismology, Part B*, pp. 1159–1173. San Diego, CA: Academic Press.

Zoback MD and Zinke JC (2002) Production-induced normal faulting in the Valhall and Ekofisk oil fields. *Pure and Applied Geophysics* 159: 403–420.

Zoback MD, Hickman S, and Ellsworth W (2010) Scientific drilling into the San Andreas fault zone—An overview of SAFOD's first five years. *Scientific Drilling* 11: 14–28.

Zoback MD, Zoback ML, Mount VS, et al. (1987) New evidence on the state of stress of the San Andreas fault system. *Science* 238: 1105–1111.

4.13 Deep Earthquakes

H Houston, University of Washington, Seattle, WA, USA

© 2015 Elsevier B.V. All rights reserved.

4.13.1	**Introduction**	329
4.13.2	**Deep Earthquake Source Properties**	330
4.13.2.1	Nature of the Seismic Source	330
4.13.2.2	Numbers and Magnitudes of Deep Earthquakes	331
4.13.2.2.1	Gutenberg–Richter statistics	331
4.13.2.2.2	Isolated giant earthquakes	331
4.13.2.3	Aftershock Occurrence	332
4.13.2.3.1	Long-range triggering	333
4.13.2.4	Rupture Processes of Deep Earthquakes	334
4.13.2.4.1	Duration of rupture	334
4.13.2.4.2	Complexity of rupture	334
4.13.2.4.3	Spatial pattern of rupture, rupture velocity, and stress drop	337
4.13.2.4.4	Radiated seismic energy	338
4.13.2.4.5	The *M*8.2 1994 Bolivia earthquake	338
4.13.2.4.6	The *M*8.3 2013 Sea of Okhotsk earthquake	339
4.13.3	**Possible Mechanisms of Deep Earthquake Generation**	339
4.13.3.1	Dehydration Embrittlement	339
4.13.3.2	Anticrack Faulting in a Metastable Phase	340
4.13.3.3	Thermal Shear Instability	342
4.13.4	**Geophysical Setting of Subducting Slabs**	343
4.13.4.1	Thermal State of Subducting Slabs	343
4.13.4.1.1	Thermal parameter	343
4.13.4.2	Stresses and Focal Mechanisms	344
4.13.4.2.1	Fault-plane solutions and reactivated faults	345
4.13.5	**Seismic Structures in Slabs**	345
4.13.5.1	Double Seismic Zones	345
4.13.5.2	Planar Shear Zones	347
4.13.5.2.1	Repeating earthquakes	348
4.13.5.3	Outboard Earthquakes and Recumbent Slabs	348
4.13.6	**Implications for Seismogenesis**	349
4.13.6.1	Viability of the Three Main Physical Mechanisms	349
4.13.6.2	Notable Points and Speculations	350
4.13.7	**Concluding Thoughts**	351
References		351

4.13.1 Introduction

Earthquakes occur below Earth's crust to nearly 700 km depth in planar regions of seismicity termed Wadati–Benioff zones, which delineate the cold core of subducting slabs. The rates of occurrence of earthquakes deeper than 70 km, however, are several times lower than those of crustal quakes. This chapter reviews the properties and possible physical mechanisms of deep earthquakes, generally defined here as earthquakes that occur below about 70 km depth. At times, the discussion distinguishes between deep and intermediate earthquakes (setting the boundary between them at about 350 km), while other times, the specific depth ranges are indicated.

The physical processes that permit the occurrence of deep earthquakes are not well understood. The brittle–frictional processes that operate to produce shallow seismic rupture, fracture, and frictional sliding appear to be suppressed by conditions at depth. In particular, increasing pressure with depth tends to inhibit fracture and sliding, while increasing temperature promotes ductile flow (e.g., Griggs and Handin, 1960; Scholz, 2002). Classic experiments on olivine show that the shear stresses necessary to overcome the high normal stresses imposed by the overburden are far greater than can be sustained at the temperatures at depth (e.g., Brace and Kohlstedt, 1980). Thus, failure must occur at lower shear stresses by one or more other mechanisms.

In addition to posing a long-standing puzzle to earthquake seismologists, deep earthquakes pose significant seismic hazards in several regions around the globe. The hazard stems primarily from relatively shallow intermediate-depth earthquakes beneath populated continental regions, because event frequency falls off exponentially with depth below about

30 km. Regions that have suffered damage from such subcrustal earthquakes include Romania (4 March 1977 M 7.5 and depth = 90 km), Chile (25 January 1939 M 7.8 and depth = 80 km), Washington State in the United States (28 February 2001 M 6.8 and depth = 54 km), and Japan (26 May 2003 M 7.0 and depth = 61 km). See Frohlich (2006a) for a discussion of the seismology and severe destruction associated with the 1977 Romania and 1939 Chile intermediate-depth earthquakes, including death tolls of about 1500 and 28 000, respectively.

Finally, deep earthquakes have played a key role in illuminating the structure of Earth's mantle and core. This is largely because their waveforms are less complicated than those of shallow earthquakes due to the separation in time of the first-arriving body waves from later-arriving body waves reflected off the Earth's surface. In addition, direct body waves from deep earthquakes need to pass through the highly layered and laterally heterogeneous crust of the Earth beneath only the receiver, rather than beneath both the receiver and source.

Deep earthquakes remain mysterious. This overview does not attempt an exhaustive survey of the topic, which impinges on many aspects of geophysics including plate tectonics, mineral physics, mantle convection, slab and mantle structure, composition, and rheology, as well as seismology. Rather, the aim is to highlight some of the more recent work on the subject and raise several issues, some perhaps overlooked. Previous reviews that the reader may find helpful include Frohlich (1989), Green and Houston (1995), Kirby et al. (1996), Green and Marone (2002), and Frohlich (2006a), an entire book devoted to deep earthquakes.

This chapter discusses seismological properties of deep earthquakes, followed by the leading candidates for physical mechanisms proposed to generate deep earthquakes, the geophysical setting (temperatures and stresses) of deep earthquakes, and seismic structures in slabs, and finally focuses on implications for seismogenesis and key outstanding issues.

4.13.2 Deep Earthquake Source Properties

Section 4.13.1 presents salient properties of deep earthquakes. Some of these have been known for decades, while others have come into focus only recently. Particular attention is drawn to changes in source properties with depth, magnitude, or subduction zone, with an eye to whether these may provide constraints on possible deep earthquake generation mechanisms.

The bimodal nature of the distribution of earthquakes with depth has been pondered for decades. The rate of occurrence of earthquakes with depth falls roughly exponentially from near 50 to ~300 km, remains at a low level until ~500 km depth, and then shows a marked resurgence before abruptly terminating near 680 km (**Figure 1**). A similar bimodal distribution is observed individually, with some variation, in every subduction zone with deep seismicity (Helffrich and Brodholt, 1991), and provides the basis of the long-standing distinction between intermediate and deep earthquakes.

To explore the three-dimensional (3D) seismicity distributions in various subduction zones interactively, reader may visit the website listed at the end of the chapter.

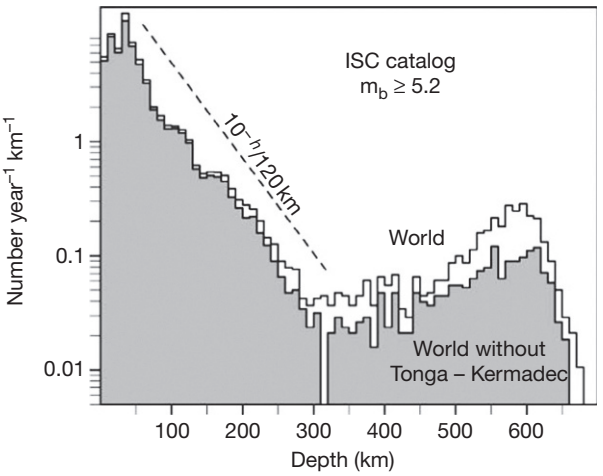


Figure 1 Earthquake occurrence rates versus depth for ISC catalog events from 1964 to 2001. The upper line corresponds to rates for the entire world and the lower line to the entire world excluding Tonga–Kermadec. The rate decreases roughly exponentially between 50 and 300 km. Adapted from Frohlich C (2006) *Deep Earthquakes*. Cambridge, UK: Cambridge University Press.

4.13.2.1 Nature of the Seismic Source

The seismic signals from deep earthquakes are described by a combination of double-couple and compensated linear vector dipole (CLVD) seismic sources. Although an implosive isotropic component of deep earthquakes has long been sought, in connection with the abrupt increases in density associated with upper-mantle phase changes, a nonzero isotropic component has not been reliably detected for a deep earthquake (Hara et al., 1996; Kawakatsu, 1996).

Careful reanalysis of low-frequency spectra of the great 1970 Columbia earthquake indicates no significant isotropic component (Dziewonski and Gilbert, 1974; Russakoff et al., 1997). Similarly, a very broadband analysis of the 1994 Bolivia earthquake limited an isotropic component to < 1–2% (Hara et al., 1995). The CLVD component, however, does appear more prevalent for deep than for shallow earthquakes, in that a higher proportion of deep earthquakes have a large CLVD component (Frohlich, 1995; **Table 1**). However, the favored interpretation of this observation is that, except for some shallow earthquakes in volcanic regions, large CLVD components of both deep and shallow earthquakes usually result from a change in focal mechanism during rupture (e.g., McGuire et al., 1997; Tibi et al., 1999), as double-couple motion on two fault planes can sum to an apparent CLVD component (Frohlich, 1994). Thus, because an isotropic component does not seem to be present and a CLVD component can be explained as double-couple motions on more than one fault plane, deep earthquakes are believed to result from shear deformation, similar to the vast majority of shallow earthquakes. However, resolution of the fault-plane geometry is insufficient to constrain the deformations to lie on a single planar surface. Thus, the possibility that a deep earthquake could consist of deformation over a wide shear zone or on multiple en echelon planes (e.g., Chen, 1995) has not been definitively ruled out.

Table 1 Largest deep earthquakes below 300 km from 1906 to 24 May 2013, ordered by seismic moment[a]

Date	*Area*	*Depth (km)*	*Moment (Nm)*	M_w
24 May 2013	Sea of Okhotsk	609	4.1×10^{21}	8.3
09 June 1994	Bolivia	647	2.6×10^{21}	8.2
31 July 1970	Colombia	623	1.4×10^{21}	8.1
17 January 1922	North Peru	664	9.4×10^{20}	7.9
17 June 1996	Flores Sea	589	7.9×10^{20}	7.9
29 March 1954	Spain	630	7.0×10^{20}	7.9
29 September 1973	North Korea	593	5.0×10^{20}	7.8
14 August 2012	Sea of Okhotsk	598	4.8×10^{20}	7.7
11 June 1972	Celebes Sea	332	4.7×10^{20}	7.7
05 July 2008	Sea of Okhotsk	611	4.5×10^{20}	7.7
19 August 2002	Fiji	699	4.3×10^{20}	7.7
26 May 1932	Fiji	560	4.0×10^{20}	7.7
15 August 1963	Peru–Bolivia	573	3.9×10^{20}	7.7
28 February 1950	Sea of Okhotsk	339	3.9×10^{20}	7.7
25 May 1907	Sea of Okhotsk	548	3.7×10^{20}	7.7
23 July 2010	Philippines	577	3.6×10^{20}	7.7
27 January 2006	Banda	397	3.5×10^{20}	7.7
09 November 1963	Western Brazil	573	3.5×10^{20}	7.7
19 August 2002	Fiji	631	3.5×10^{20}	7.7
19 August 1961	Peru–Bolivia	620	3.4×10^{20}	7.7
09 March 1994	Fiji	563	2.7×10^{20}	7.6
23 May 1956	Fiji	436	2.7×10^{20}	7.6
26 July 1958	Peru–Bolivia	592	2.6×10^{20}	7.6
08 August 2007	Java	305	2.6×10^{20}	7.6

[a]Updated from Table 1.2 of Frohlich (2006a), which was based on Huang and Okal (1998) augmented by the CMT catalog for events occurring since 1996.

4.13.2.2 Numbers and Magnitudes of Deep Earthquakes

Less than a fifth of the Earth's seismicity occurs below 100 km (**Figure 1**). Eighteen percent of the ~6000 earthquakes with moment magnitude $>M_w$ 5.8 in the Harvard centroid moment tensor (CMT) catalog from 1977 to 2006 have CMT depths >100 km, and only 6% occur at depths between 400 and 700 km, below which no earthquakes have been reliably located (see Frohlich (2006a) for a discussion). Only about 2.2% occur within the seismically quiet interval from 300 to 500 km depth.

On the other hand, the vast majority of Earth's intraplate (within a tectonic plate) seismicity occurs below 50 km as deep earthquakes within subducted slabs (particularly so if continental seismicity near diffuse plate boundaries is excluded). Clearly, oceanic lithosphere becomes more seismogenic after it has subducted beneath Earth's surface. Thus, it appears that one or more of the following processes associated with subduction have seismically activated the oceanic lithosphere – the damage and/or hydration from slab bending at the trench and outer rise, the high temperatures and pressures, or the high internal stresses imposed on subducting slabs by negative buoyancy, thermal expansion, and volumetric phase changes. These phenomena are discussed in **Sections 4.13.4** and **4.13.5.1**.

4.13.2.2.1 Gutenberg–Richter statistics

Like shallow earthquakes, deep earthquakes generally follow Gutenberg–Richter statistics; that is, the number of earthquakes with magnitude greater than M is given by $n = aM^b$ where a and b are constants. For global sets of both deep and shallow earthquakes, the value of b is near 1.0 (Frohlich and Davis, 1993; Kagan, 1999). However, although regional variations are, in general, weak for shallow earthquakes, a remarkable systematic regional variation in the values of b is found for the deepest events (e.g., Frohlich, 1989, Table 1, 2006a; Frohlich and Davis, 1993; Giardini, 1988). For events below 400 km, Giardini and Lundgren (1995) obtained b-values that varied from ~0.4 for South America to ~1.2 for Tonga (**Figure 2**). Although the Tonga subduction zone lays claim to the largest number of deep earthquakes, South America hosts the largest-magnitude deep ruptures. While Tonga has about ten times more M 5.5 events, the two regions have similar numbers of M 6.5 events, and the biggest deep earthquakes occur in South America (**Figure 2**). With respect to inferred slab temperatures (discussed in **Section 4.13.4.1**), Tonga and South America are cold and warm end-members, respectively, of subducted slabs with deep seismicity. Strong thermal control of the deep magnitude–frequency relation is, therefore, implicated (e.g., Wiens, 2001; Wiens and Gilbert, 1996). Interestingly, the b-values of seismicity between 100 and 350 km deep are not significantly different for Tonga and South America. Evidently, thermal influence on earthquake generation does not operate in the same way above 350 km.

4.13.2.2.2 Isolated giant earthquakes

At depths >400 km, the very largest earthquakes tend to occur in South America, in isolated settings, and produce very few aftershocks for their size. A remarkable disparity, one related to the regional contrast in b-value noted earlier, is that the three largest-known earthquakes deeper than 300 km occurred in South America, whereas only 10% of M_w 6.0 earthquakes

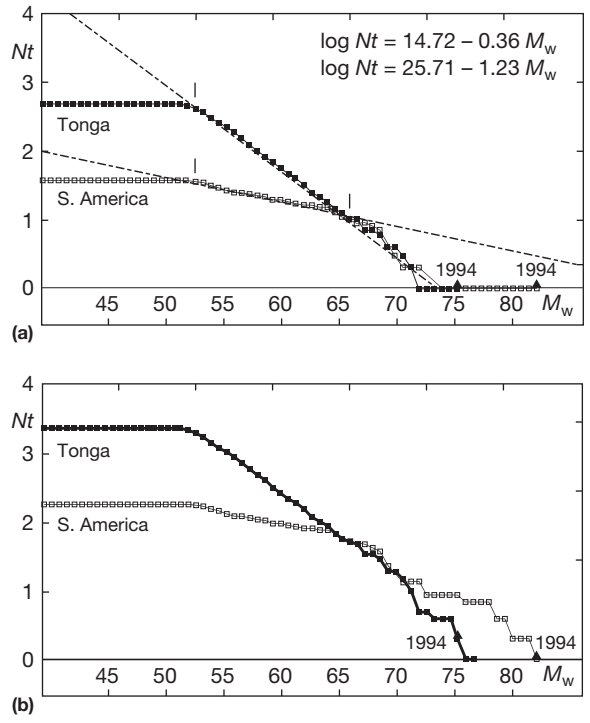


Figure 2 Magnitude–frequency relations for earthquakes below 400 km depth in Tonga (solid squares) and South America (open squares). The logarithm of the number of events with magnitude greater than a given value is plotted. (a) From Harvard CMT catalog from 1977 to 1994. Dashed lines, for which equations are given in the upper right, show fits to the data for magnitudes between M_W 5.3 and 6.6. Very different *b*-values (slopes) for the two regions were determined. The 1994 Bolivia and Tonga earthquakes are roughly consistent with the magnitude–frequency behavior at lower magnitudes. (b) From hybrid catalog extended back to 1900 by Giardini and Lundgren (1995) by including magnitudes from the literature for larger events and assuming that the magnitude–frequency behavior (of events up to M_W 7.2) during 1977–94 is typical of that from 1900 to 1994. Adapted from Giardini D and Lundgren P (1995) The June 9 Bolivia and March 9 Fiji deep earthquakes of 1994: II. Geodynamic implications. *Geophysical Research Letters* 22: 2281–2284.

occur there compared to 60% in Tonga (from the Harvard CMT catalog between 525 and 650 km depth). Furthermore, 7 out of 19 of the largest earthquakes below 300 km occurred in South America compared to only five in Tonga–Fiji (see **Table 1**). The fifth largest deep earthquake, the 1954 *M* 7.9 Spanish earthquake, also occurred in an isolated setting, in lithosphere associated with a previous episode of subduction (Blanco and Spakman, 1993).

There are several peculiarities of the 'South American giants' (e.g., Estabrook, 1999). The 1994 *M* 8.2 Bolivia (**Section 4.13.2.4.5**), 1970 M_w 8.1 Colombia, and 1922 M_w 7.9 North Peru earthquakes are the three largest-known deep earthquakes below 300 km (**Table 1**). They occurred in seismically very quiet to relatively quiet regions that lie one to several hundred kilometers from well-defined Wadati–Benioff zones. Body-wave studies of the rupture directivity of the 1994 and 1970 events, as well as the 1963 M_w 7.7 Peru–Bolivia earthquake that occurred about 200 km west of the 1994 event, indicate that their rupture velocities are slower than typical, about 0.2–0.4 of the shear-wave velocity (Estabrook, 1999; Furumoto, 1977). Their seismic efficiencies (**Sections 4.13.2.4.4** and **4.13.2.4.5**) are correspondingly low as well (Estabrook, 1999). Source time functions of these three South American earthquakes (1994 Bolivia, 1970 Colombia, and 1963 Peru–Bolivia) begin with a small-to-moderate subevent followed by the main moment release 10–20 s later (Brustle and Muller, 1987; Estabrook, 1999; Fukao and Kikuchi, 1987; Furumoto, 1977). The 1963 M_w 7.7 Western Brazil earthquake (**Table 1**) did not have such a precursor in its time function, although it did have a low rupture velocity (Fukao, 1972; Fukao and Kikuchi, 1987).

The contrasts between seismicity in the Tonga and South American slabs may indicate different physical mechanisms for Tongan and South American deep earthquakes (e.g., Wiens and McGuire, 1995), or perhaps implicate different mechanisms for nucleation versus propagation of deep rupture. It appears that earthquakes can nucleate more easily in Tonga but propagate into larger ruptures more easily in South America. Thus, the properties that allow rupture to initiate appear to differ from those that promote its propagation. But further, the disparity between the contrast in *b*-values below ~350 km and the similarity of *b*-values above ~350 km suggests a difference between rupture mechanisms above and below that depth. Possible mechanisms are discussed in **Sections 4.13.3** and **4.13.6**.

4.13.2.3 Aftershock Occurrence

Deep earthquakes are notable for producing very few aftershocks (Frohlich, 1987), although there are a few cases (e.g., 1994 Tonga) in which the aftershock sequences of deep earthquakes are as numerous as those for shallow earthquakes (Wiens and McGuire, 2000). In general, the incidence of aftershocks for mainshocks with depths ≥50 km is an order of magnitude lower than for shallow mainshocks (Frohlich, 1987). Nevertheless, the aftershocks produced appear to follow, on average, Omori's law of aftershock decay (Nyffenegger and Frohlich, 2000).

Perhaps due to the paucity of deep aftershocks, Frohlich (1987) found little difference in deep aftershock production by geographic region. Subsequently, Wiens and Gilbert (1996) investigated regional variations in aftershock production for large mainshocks below 400 km and found that the most productive aftershock sequences are limited to cooler slabs, as defined by the thermal parameter (**Section 4.13.4.1.1**). Mainshocks in slabs with smaller thermal parameters, such as South America, Japan, and Izu Bonin, exhibited lower aftershock productivity than mainshocks in Tonga, Indonesia, and Marianas, which have larger thermal parameters. A more recent study of 290 somewhat smaller mainshocks failed to detect a clear relation of thermal parameter to aftershock productivity for either intermediate or deep earthquakes (Persh and Houston, 2004b). Nevertheless, for aftershocks deeper than 400 km, the situation probably parallels that of *b*-values of seismicity in general as discussed earlier, that is, *b*-values and aftershock productivity are high (i.e., more small events relative to large ones) for cooler slabs and low for warmer slabs.

Strong depth dependence of aftershock productivity is seen in **Figure 3** from Persh and Houston (2004b); their study

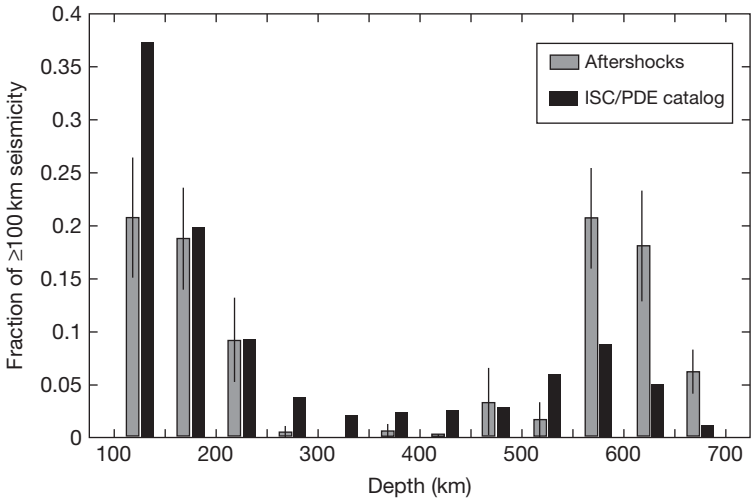


Figure 3 Depth distribution of aftershocks (light shading) compared to all deep seismicity (dark shading) from the ISC/PDE catalog. Aftershock numbers were normalized to account for mainshock moment (Persh and Houston, 2004b). The heights of the bars of each type sum to 1.0. The error bars on the aftershocks are one standard deviation and were determined by bootstrap resampling of the mainshocks in each depth bin. The distribution of aftershocks with depth differs from that of the background seismicity; at 300–550 km, aftershock productivity relative to all seismicity is much weaker than at intermediate depths, whereas deeper than 550 km, relative productivity is much stronger than at intermediate depths. Adapted from Persh SE and Houston H (2004) Strongly depth-dependent aftershock production in deep earthquakes. *Bulletin of the Seismological Society of America* 94: 1808–1816.

analyzed 290 mainshocks with $M_w \geq 6.3$ below 100 km from 1977 to 2004. Between 300 and 550 km depth, aftershock productivity falls to one-third of the relatively low level that prevails between 100 and 300 km compared to shallow crustal events. But below 550 km, productivity rises abruptly to twice the level between 100 and 300 km. **Figure 4(a)** shows the number of aftershocks per mainshock normalized to an M_w 8.2 mainshock (the largest in the data set), assuming the number of aftershocks scales linearly with $M_0^{2/3}$. Bootstrap analyses of the data indicate that the differences in aftershock productivity are statistically significant. The strong dependence of aftershock productivity on depth suggests that the reason few earthquakes occur in the depth range 300–550 km may be due to a difficulty in nucleating ruptures rather than due to a lack of ambient deviatoric stress in that depth range, which is the conventional explanation and implied by the simple thermomechanical stress model discussed in **Section 4.13.4.2**. A mainshock of a given size between 300 and 550 km presumably produces a stress change similar in magnitude to that produced by a shallower or deeper event. The abrupt increase in aftershock productivity below 550 km is striking and perhaps indicative of a change in physical mechanism near that depth.

4.13.2.3.1 Long-range triggering

Large deep earthquakes can trigger other large deep quakes at great distances. Such long-range triggering is known to occur for shallow crustal earthquakes (e.g., Hill et al., 1993). In one notable case in 2002, a M_w 7.6 event deep in the Tonga subduction zone (598 km deep) likely triggered a M_w 7.7 event (664 km deep) 313 km away 7 min later (Tibi et al., 2003b). The latter event occurred well below the bottom of the previously seismically active zone in the region. Tibi et al. (2003b) gave several other examples of triggered deep seismicity below 450 km, for which the time delays range from 2 to 143 min and the spatial separations range from 70 to 320 km. Such long-range triggering occurs mainly in Tonga but is also seen in Bolivia and Japan. For example, just 10 min after the 1994 M_w 8.2 Bolivia earthquake, an M_w 6 event was triggered about 330 km east of the mainshock at 671 km depth (Myers et al., 1995).

The triggered events were deeper than the mainshocks except in one case and often occurred in nearly aseismic regions next to or below the seismically active zone (Tibi et al., 2003b). The time delay of the triggered events does not correspond to the arrival of large seismic phases from the triggering events, suggesting that the triggered events take a few minutes to several tens of minutes to nucleate following the stress disturbance associated with the passage of the dynamic waves. The time delay could be related to the development of plastic instabilities (Tibi et al., 2003b) or to the growth of anticracks in a metastable phase, to single out two possible mechanisms (**Sections 4.13.3.2** and **4.13.3.3**). The triggered seismicity is located, presumably, in regions on the verge of failure. The location of the triggered events, near the edges of the seismically active slab and/or deeper than the triggering events, seems more suggestive of transformational faulting because compared to conditions at the triggering event, the edge of the slab is a region of less kinetic hindrance and the overstep of a metastable phase transition (i.e., heat release and Gibbs free energy) increases with depth, both of which would favor transformation. In contrast, at the relevant depths (~600 km), the instability parameter for thermal shear instability typically decreases with depth and thus is smaller near the triggered event than near the triggering event so that shear instability would be less favored at the site of the triggered event than at the site of the triggering event (**Section 4.13.3.3** and **Figure 13**).

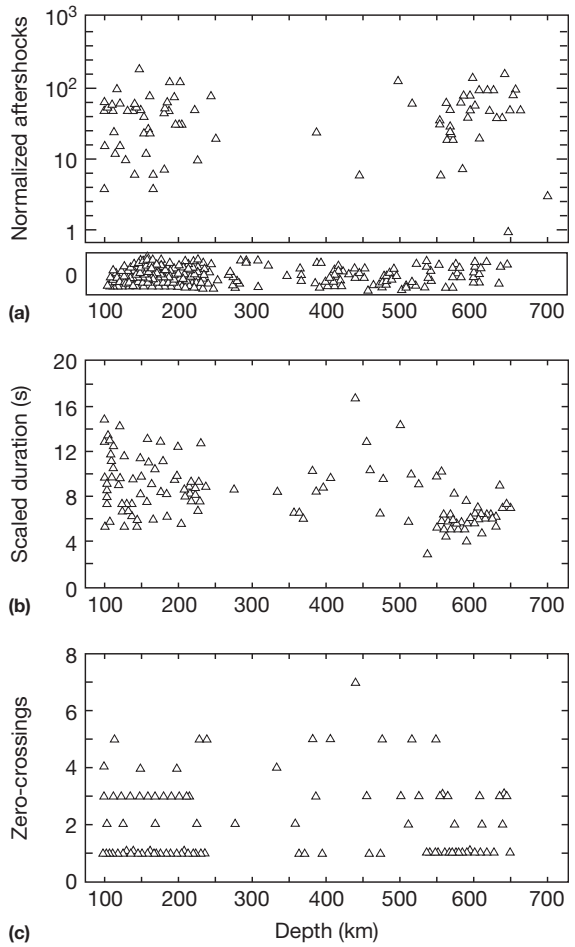


Figure 4 Depth dependence of three source properties. Note differences in properties for three depth ranges (100 to ~350 km, ~350–550 km, 550–700 km), as well as abrupt change of properties near 550 km. (a) Normalized aftershocks. The number of aftershocks was scaled to remove effect of mainshock size. There are almost no aftershocks in the middle depth range, and a larger fraction of mainshocks below 550 km have aftershocks than mainshocks in the first depth range (Persh and Houston, 2004b). Events with no aftershocks are plotted below with random jitter added to reduce overlap. (b) Scaled duration. Durations determined from source time functions were scaled to remove the effect of earthquake size (Persh and Houston, 2004a). (c) Source complexity measured by zero-crossings of the time derivative of low-passed, scaled source time functions (Persh and Houston, 2004a). Complexity appears greater in the middle depth range. Adapted from Persh SE and Houston H (2004) Strongly depth-dependent aftershock production in deep earthquakes. *Bulletin of the Seismological Society of America* 94: 1808–1816.

4.13.2.4 Rupture Processes of Deep Earthquakes

As with shallow earthquakes, our knowledge of the rupture process of deep earthquakes (i.e., the progression of rupture through space and time) comes from analysis of the shape and duration of seismic waveforms, particularly the direct P and S body waves. Unlike for shallow earthquakes, for deep earthquakes, these waves are not complicated by the depth-reflected phases pP, sP, and sS, because those phases arrive long after the direct wave. In addition, the paths of body waves from deep earthquakes traverse a smaller length of the highly heterogeneous material near Earth's surface and, therefore, suffer less attenuation, both intrinsic and scattering. Thus, deep earthquake body waves are relatively clean, simple, and sharp (**Figure 5**) compared to the body waves from shallow earthquakes. In the analysis of deep earthquake source properties, it is advisable to use depth-dependent values for such parameters as seismic velocity, rigidity, attenuation (e.g., t^*), and geometric spreading (which is also distance dependent), as these vary significantly from the top to the base of the upper mantle.

4.13.2.4.1 Duration of rupture

It has long been clear that deep earthquakes (especially very deep ones) tend to be briefer than shallow ones of similar magnitude. It is relatively direct to obtain the duration of rupture from broadband seismograms, although care must be taken to determine the time of termination (**Figure 5**). Although there is significant variability, earthquake durations generally scale as seismic moment to the one-third (e.g., Houston, 2001; Houston et al., 1998; Kanamori and Anderson, 1975; Vidale and Houston, 1993). Thus, to determine how duration may depend on various parameters, it is helpful to consider scaled duration, which is duration divided by moment (relative to a reference moment) raised to the one-third power. Scaled durations from a global set of time functions of shallow and deep earthquakes generally decrease gradually with increasing depth, but abrupt decreases are seen near 40 and 550 km (Houston, 2001). The abrupt decrease near 40 km is interpreted as due to the contrast in stress drop and/or shear velocity between ruptures occurring between plates (the global data set used consisted mostly of thrust earthquakes in subduction zones) and those within plates (below 40 km). The abrupt decrease near 550 km was corroborated by Persh and Houston (2004a) using stacks of carefully aligned broadband P waves of 111 large events below 100 km (**Figures 4–6**). Tocheport et al. (2006) obtained similar results using a similar data set and a more automated alignment and stacking method. The durations deviate somewhat from a simple model in which rupture velocity is a constant fraction of shear velocity and all else (e.g., stress drop and shape of rupture area) remains constant (see Persh and Houston (2004a); **Figure 6**). Such a deviation near 550 km, within considerable variability, can also be detected in Frohlich (2006a)'s Figure 6.4, although that author prefers an interpretation involving no deviation from the simple model. As with the aftershock productivity discussed earlier (**Section 4.13.2.3**) and the rupture complexity discussed next, the abrupt change in scaled duration near 550 km suggests the possibility of a change in the physical mechanism of rupture.

4.13.2.4.2 Complexity of rupture

As in shallow earthquakes, deep earthquake ruptures are often complex, with multiple rupture subevents (e.g., **Figure 6**). A significant fraction, roughly half, of events below 100 km depth exhibit multiple subevents in their waveforms and time functions (e.g., Antolik et al., 1999; Fukao and Kikuchi, 1987; Houston and Vidale, 1994; Houston et al., 1998; Persh and Houston, 2004a; Wu and Chen, 2001; Tibi et al., 2002, 2003a). Houston et al. (1998) proposed that complexity be measured

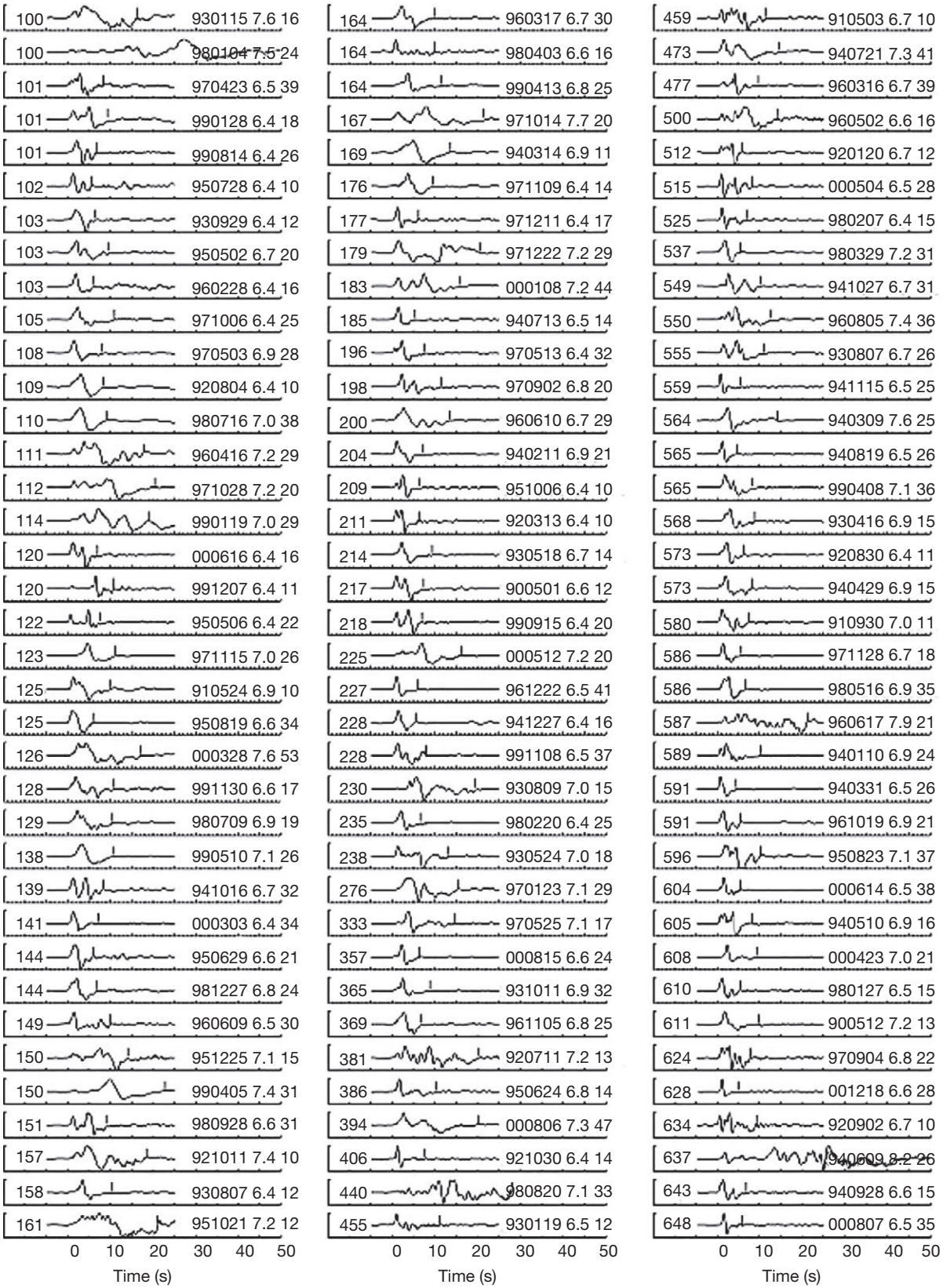


Figure 5 Stacks of P-wave velocity seismograms for 111 large earthquakes deeper than 100 km. Stacks were constructed by aligning P waves on a small feature early in the rupture, then summing seismograms. Depth, date, M_W, and number of stations stacked are given for each event. Events are ordered by depth. Adapted from Persh SE and Houston H (2004) Deep earthquake rupture histories determined by global stacking of broadband P waveforms. *Journal of Geophysical Research* 109. http://dx.doi.org/10.1029/2003JB002762.

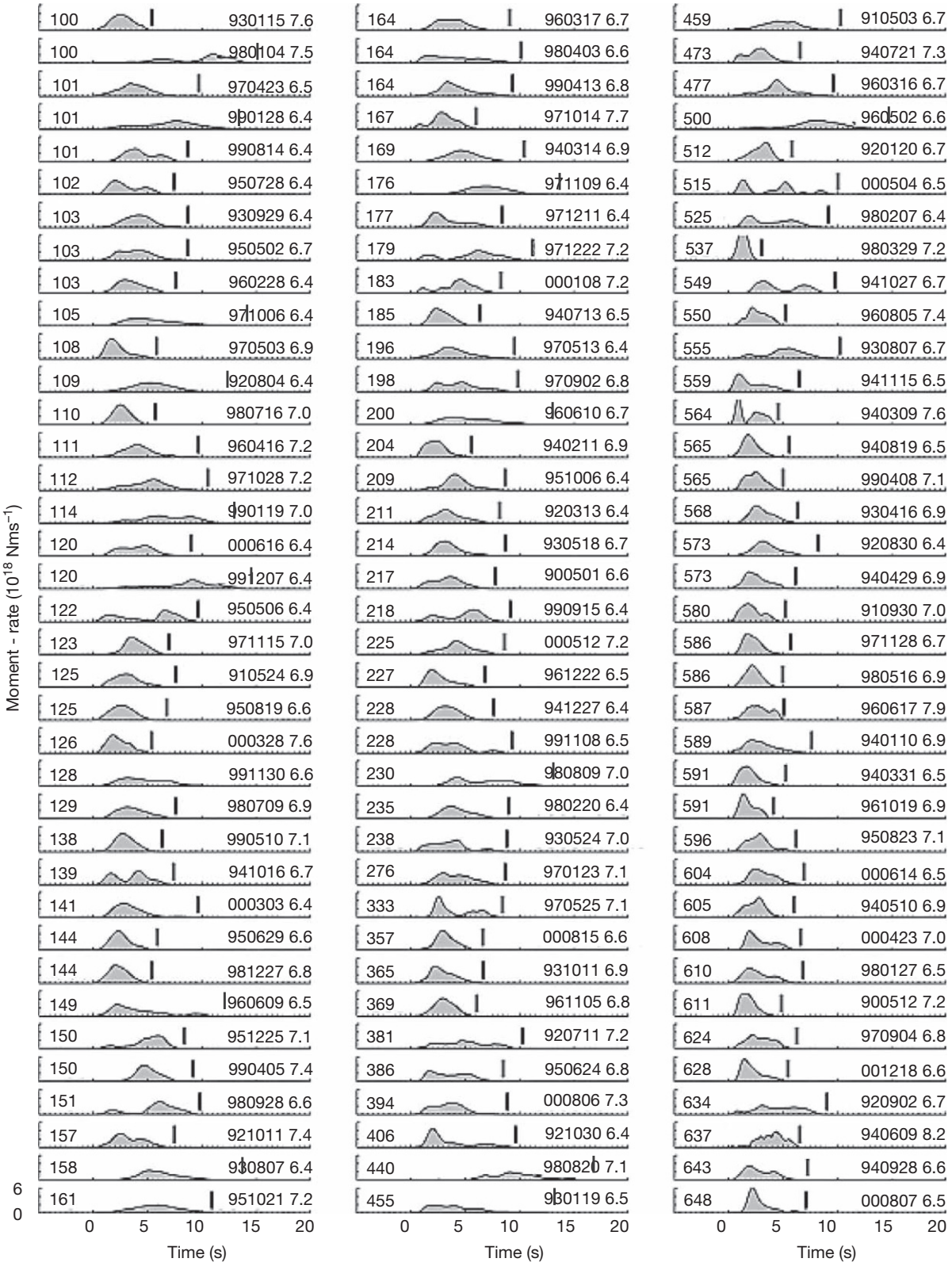


Figure 6 Scaled source time functions in displacement from P-wave stacks in **Figure 5**. Time functions were scaled to remove effect of earthquake size. Time and amplitude are scaled by the cube root and the squared cube root, respectively, of the seismic moment relative to a reference moment 10^{19} Nm. Depth, date, and M_W are shown for each event. Events are ordered by depth. Note simpler shapes for time functions deeper than ~550 km. Adapted from Persh SE and Houston H (2004) Deep earthquake rupture histories determined by global stacking of broadband P waveforms. *Journal of Geophysical Research* 109. http://dx.doi.org/10.1029/2003JB002762.

quantitatively by the number of zero-crossings of the time derivative of the scaled source time function, after low-pass filtering to equalize the effect of event magnitude. Using this approach, Houston et al. (1998) and Persh and Houston (2004a) found that events between 350 and 550 km depth have greater complexity (**Figures 4** and **6**) compared to events between 100 and 350 km, in addition to being longer than expected, while events below 550 km are simpler than either group above them, in addition to being briefer as discussed in **Section 4.13.2.4.1**. Thus, in the 350–550 km depth range, events are more complex and longer in scaled duration relative to shear velocity and have fewer aftershocks, compared to events above and below that depth range (**Figure 4**). Perhaps the aftershocks of events in the 350–550 km depth range occur immediately as part of the main rupture, leading to fewer measured aftershocks but greater complexity of the mainshock. The abruptness of these changes in several attributes (i.e., complexity, aftershock productivity, and duration) at about 550 km depth may suggest a change in earthquake generation mechanism (Persh and Houston, 2004a,b).

4.13.2.4.3 Spatial pattern of rupture, rupture velocity, and stress drop

The progression of rupture in space and time is a key characteristic that bears on the mechanics and physics of fault failure. The variation with azimuth of body waveforms provides the main constraint on the progression of rupture (**Figure 7**). Under most circumstances, what can be resolved is the velocity of propagation of a significant amount of slip, rather than the propagation of a crack tip. Many studies have forward-modeled or inverted P and SH waveforms based on parameterizations of deep ruptures as line sources or multiple point sources (e.g., Glennon and Chen, 1995; Tibi et al., 2002, 2003a). Such parameterizations yield some aspects of the spatiotemporal distribution of rupture, including fault length and rupture velocity. For example, Glennon and Chen (1995) and Tibi et al. (2002) modeled eight and six events, respectively, with a line source, obtaining a length of rupture, source time functions, and average rupture velocities. Although the 2D pattern of slip over a fault plane is more difficult to constrain, a number of studies have determined 2D slip in large deep earthquakes (e.g., Antolik et al., 1999; Fukao and Kikuchi, 1987; McGuire et al., 1997; Wu and Chen, 2001, as well as numerous studies of the 1994 Bolivia earthquake summarized in **Section 4.13.2.4.5**).

Determination of the spatial distribution of slip in deep earthquakes is difficult due to several factors. Deep earthquakes tend to be far more compact for a given moment than shallow ones. There is rarely an abundance of well-distributed low-noise seismometers. And because few or no aftershocks occur to delineate the rupture surface, it is usually difficult to identify which of the nodal planes is the fault plane, although Warren et al. (2007), Warren et al. (2008), and Myhill and Warren (2012) developed and applied an approach to identify which of the two nodal planes hosted rupture using waveform shape to resolve directivity. There appears to be a tendency for ruptures, particularly those of intermediate-depth earthquakes, to propagate along strike and near-horizontally (Antolik et al., 1999; Kiser et al., 2011; Myhill and Warren, 2012; Tibi et al., 2002; Warren et al., 2007, 2008).

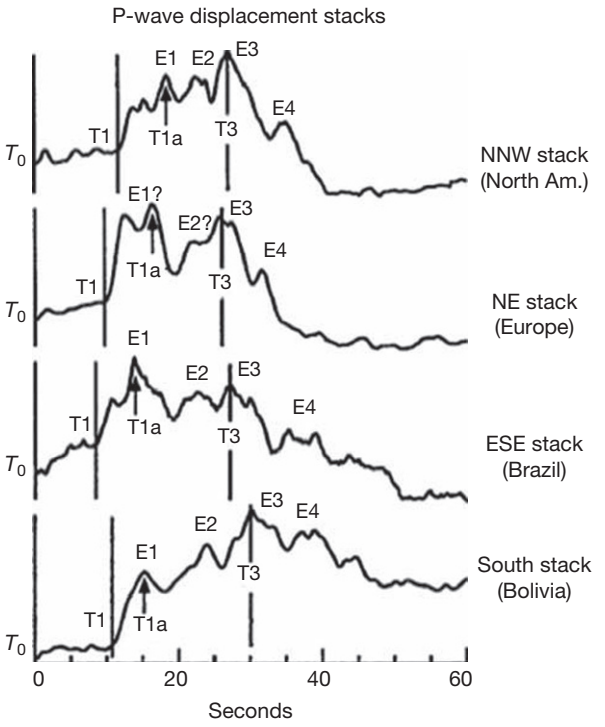


Figure 7 Stacked P waveforms of the 1994 Bolivia earthquake in different azimuthal directions. The variations with azimuth constrain the location and timing of high-slip regions. Labels indicate timing of subevents within the rupture. The earthquake initiated at time T_0, but moment-release rate increased sharply about 11 s later. Adapted from Beck SL, Silver P, Wallace TC, and James D (1995) Directivity analysis of the deep Bolivia earthquake of June 9, 1994. *Geophysical Research Letters* 22: 2257–2260.

Taken together, the results of existing studies indicate that there is much variability in deep earthquake fault geometry and aspect ratios, rupture velocities, and stress drops (e.g., see Frohlich, 2006a).

Rupture velocities are typically given as a fraction of shear-wave velocity β in the source region. The distribution of ~100 estimates of rupture velocities from various studies is shown in **Figure 8** after Frohlich (2006a). Notably, rupture velocities are lower on average and vary more for deep than for shallow earthquakes – from 0.2β to 0.9β (Frohlich, 2006a) or, arguably, even faster (e.g., Kuge, 1994). For most intermediate and deep earthquakes, rupture velocities range from 0.4β to 0.9β (e.g., Suzuki and Yagi, 2011 and Tibi et al., 2002, 2003a; Warren et al., 2007), but they are notably slower (0.2–0.4β) for very large deep South American events – 1963 Peru–Bolivia, 1970 Colombia, and 1994 Bolivia (Beck et al., 1995; Estabrook, 1999; Furumoto, 1977). A general dependence of rupture velocity on slab thermal state has been proposed, in which earthquakes in warmer slabs rupture more slowly (Tibi et al., 2003a; Wiens, 2001). The variability and difficulty of resolution of rupture velocity currently preclude a definitive evaluation of that proposal.

Static stress drops are difficult to determine accurately. Based on detailed waveform inversion studies, they range more than an order of magnitude from ~1 to ~50 MPa (e.g., Antolik et al., 1999; Tibi et al., 2002, 2003a) and even higher for the Bolivia

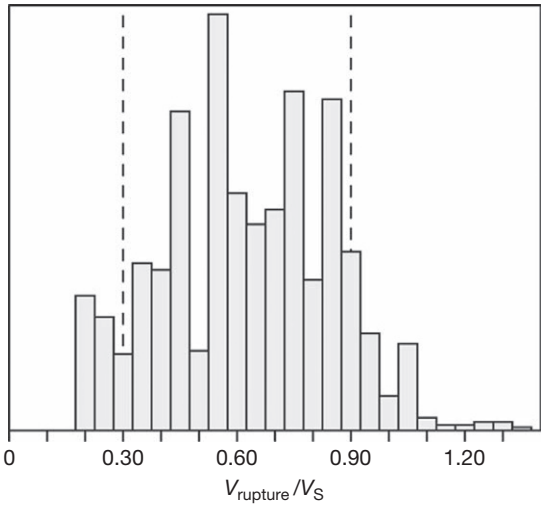


Figure 8 Histogram of reported rupture velocities from a compilation of source studies of events with $M_W \geq 6.3$ and depth between 60 and 700 km. Velocities are plotted as a fraction of shear velocity at the source depth. About 80% of the estimates fall between the dashed lines. Adapted from Frohlich C (2006) *Deep Earthquakes*. Cambridge, UK: Cambridge University Press.

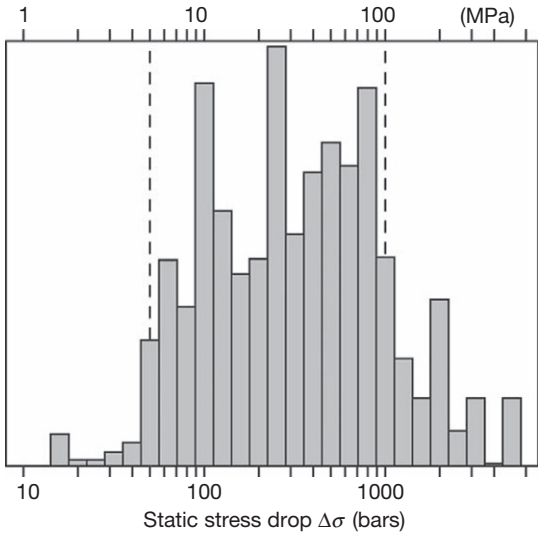


Figure 9 Histogram of reported static stress drops from a compilation of source studies of events with $M_W \geq 6.3$ and depth between 60 and 700 km. About 80% of the estimates fall between the dashed lines. Adapted from Frohlich C (2006) *Deep Earthquakes*. Cambridge, UK: Cambridge University Press.

event, discussed later. **Figure 9** taken from Frohlich (2006a) shows ~100 estimates of static stress drops of deep earthquakes compiled from numerous studies. As Antolik et al. (1999) pointed out, difficulties in resolving details of slip due to attenuation at teleseismic distances and limited azimuthal station coverage result in an inferred spatial slip distribution that is smoothed and broadened compared to the actual slip, so that actual static stress drop over the fault plane is likely more heterogeneous and, in particular, larger than many studies infer. Static stress drops of deep earthquakes average about an order of magnitude larger than those of shallow events, and thus, the strain drops of deep earthquakes average about three times higher than in shallow events due to the increase in rigidity.

4.13.2.4.4 Radiated seismic energy

Seismically radiated energy is a fundamental measure of an earthquake source and can be determined from P and S body waves. Whereas longer-period waves are suitable for measuring seismic moment, seismic energy is best measured by broadband waves that contain the corner frequency, which is the inverse of the rupture duration. Therefore, radiated energy is typically determined from body waveforms. It is proportional to the integral over time (or frequency) of squared far-field velocity. Houston and Williams (1991) calculated radiated energies from broadband records and used them to obtain Orowan stresses, which are twice the apparent stresses (the rigidity times the energy/moment ratio), for 68 globally distributed earthquakes deeper than 100 km. Values of Orowan stresses and apparent stresses are three to five times larger than those for most shallow earthquakes (Choy et al., 2006), which occur in plate boundaries (interplate), but are comparable to those of outer-rise earthquakes, which occur within the oceanic lithosphere (Green and Houston, 1995). No robust dependence on depth was detected (Houston and Williams, 1991).

Radiation efficiency is the ratio of radiated energy to radiated plus fracture energy (see, e.g., Venkataraman and Kanamori, 2004) and is equal to twice the ratio of apparent stress to static stress drop. High values (i.e., near 1) imply that much of the available nonfrictional energy change in the earthquake went into the seismic radiation; low values imply that a high proportion of the energy was dissipated in a 'fracture' process. While the energy/moment ratios of deep earthquakes are, perhaps surprisingly, not dissimilar to those of shallow quakes, some large deep events have very low radiation efficiency due mainly to inferred high static stress drops. A few large events below 400 km have very low values of radiation efficiency (<0.1) (Tibi et al., 2003a; Venkataraman and Kanamori, 2004; Winslow and Ruff, 1999); these events occurred in warmer-to-moderate slabs (of slabs possessing deep events) with somewhat lower thermal parameters (Estabrook, 1999; Tibi et al., 2003a). The events include the 1994 Bolivia, 1970 Colombia, 1963 Peru–Bolivia, 1994 Japan Sea, 1999 Eastern China, and 2002 Eastern China earthquakes. This has been interpreted as evidence for energy dissipation on the fault zone through a thermal process such as melting during rupture (e.g., Kanamori et al., 1998; Tibi et al., 2003a; Venkataraman and Kanamori, 2004).

4.13.2.4.5 The M8.2 1994 Bolivia earthquake

Many studies have examined the spatiotemporal distribution of slip in the second largest-known deep earthquake, the 1994 M_w 8.2 Bolivia earthquake (e.g., Chen, 1995; Estabrook and Bock, 1995; Ihmle, 1998; Kikuchi and Kanamori, 1994; Lundgren and Giardini, 1995; Silver et al., 1995). The data sets used for these studies were largely similar suites of body waveforms, although somewhat different methods of inversion were applied. The results consistently locate the highest slip ~20 km northeast of the hypocenter occurring ~20 s after seismic rupture initiation. The spatial distribution implies a high static stress drop of ~100 MPa (e.g., Kikuchi and Kanamori, 1994). Because the high slip is located northeast of the hypocenter, but at about the same depth, the near-horizontal nodal plane of the focal mechanism is generally

thought to be the rupture plane, and a slow rupture velocity of 1–2 km s^{-1} is inferred. This is roughly consistent with the aftershock distribution (Myers et al., 1995). However, Chen (1995) fit the waveforms with relatively fast-growing circular ruptures on three closely spaced near-vertical planes in an en echelon arrangement. In that study, the centroids of the large vertical planes are located near the high-slip region found in the other studies, so that the apparent rupture speed is again 1–2 km s^{-1}. Chen's model has not been ruled out, which illustrates the difficulty of resolving detailed patterns of slip even in large deep earthquakes (this difficulty still persists for shallow earthquakes in which the fault plane is well known). Thus, the argument that the rupture plane of the Bolivia earthquake is too wide to fit within a metastable olivine wedge can perhaps be circumvented (**Section 4.13.6.1**). In this regard, thickening of the slab above 670 km due to resistance to penetration into the lower mantle has also been suggested (Kirby et al., 1996). However, most thermokinetic modeling to date argues against the persistence of metastable olivine to the base of the upper mantle in any but the coolest slabs and particularly not in warmer ones such as South America (Devaux et al., 1997; Mosenfelder et al., 2001; see **Section 4.13.4.1**).

The shape of the 1994 Bolivia earthquake source time function is quite distinctive (**Figures 6** and **7**). After an initial rapid rise of ~2 s, a relatively low level of moment release (corresponding to an M_W 7.3 event) is sustained for ~10 s, followed by a strong breakout phase of ~30 s duration consisting of multiple strong pulses (e.g., Beck et al., 1995; Kikuchi and Kanamori, 1994; Persh and Houston, 2004a). A similar delay for the major episode of moment release is seen in the 1970 Columbia and the 1963 Peru–Bolivia time functions (Brustle and Muller, 1987; Estabrook, 1999; Fukao and Kikuchi, 1987; Furumoto, 1977). Such a time function shape suggests that one physical mechanism of failure may occur in the initial stages of the deep South American giants followed by a different mechanism that generates most of the slip. Consideration of the space–time slip distribution, as well as the shape of the time function, suggests a difference between the locations within the slab where a rupture can initiate easily and where it can grow the largest (**Section 4.13.2.2.2**).

Kanamori et al. (1998) considered the energy balance of the Bolivia earthquake. The radiation efficiency, which is the maximum seismic efficiency, is estimated from the radiated energy-to-moment ratio and the static stress drop. A small value for radiation efficiency, 0.036, is obtained, which leads to a minimum frictional stress of 55 MPa. From the minimum frictional stress, Kanamori et al. (1998) calculated that melting would occur if the thickness of the fault zone over which the displacement takes place was <0.3 m. The very low value of radiation efficiency indicates that most of the energy released was dissipated rather than radiated as seismic waves, which is consistent with the low rupture velocities determined by most studies (e.g., Beck et al., 1995; Kikuchi and Kanamori, 1994). Kanamori et al. (1998) inferred that once slip was initiated and sufficient displacement took place, melting probably prolonged the rupture (**Section 4.13.3.3**).

4.13.2.4.6 The M8.3 2013 Sea of Okhotsk earthquake

The largest-known deep earthquake occurred in the Kamchatka–Kurile slab under the Sea of Okhotsk on 24 May 2013 at a depth of ~610 km (global centroid moment tensor (GCMT) catalog). The GCMT focal mechanism is consistent with down-dip compression, although the slab geometry is poorly constrained due to limited background seismicity (Zhan et al., 2014). Eight aftershocks with M 4.1–4.5 occurred, as well as a M_W 6.7 event ~200 km south of the region of largest slip (Ye et al., 2013). Four months later, another M_W 6.7 occurred about ~70 km south of the largest slip. Large earthquakes in the general area precede the event as well (**Table 1**). The aftershock and background activity are notably more vigorous than that of the 1994 Bolivia great earthquake, consistent with higher aftershock productivity and b-values in cooler slabs (**Section 4.13.2.3**).

Initial body-waveform inversions (Ye et al., 2013; Zhan et al, 2014) estimate a rupture dimension of 90 km (Zhan et al., 2014) to 180 km (Ye et al., 2013) along a direction subparallel to the strike of the nodal planes and that inferred for the slab. They also yield a stress drop of ~4 MPa and a rupture velocity of ~4 km s^{-1}. Rupture on the subhorizontal plane is weakly favored. Considerations of radiation efficiency constrain the width of the rupture to $\lesssim 60$ km and thus imply a higher static stress drop (~14 MPa), as well as a radiated energy-to-moment ratio about twice the 1994 Bolivia event (Ye et al., 2013). These results suggest a less-dissipative source character that contrasts with that of the Bolivia event, consistent with the different thermal states of the South American and Kamchatka–Kurile slabs. Considering that the geometry of the slab is poorly known, these results may be consistent with transformational faulting in metastable olivine driving the 2013 Sea of Okhotsk event and initiating the 1994 Bolivia event, which then developed thermal shear failure progressing to melting (Kanamori et al., 1998; Zhan et al., 2014).

4.13.3 Possible Mechanisms of Deep Earthquake Generation

Here, the author outlines the three primary mechanisms proposed for the generation of deep earthquakes. These mechanisms have some distinct implications for deep earthquake locations and source processes, but data are as yet insufficient to distinguish definitively between them. The mineralogy of subducting slabs is sufficiently complex that a number of somewhat more specific mechanisms have been proposed over the years, but limitations on space preclude their discussion.

4.13.3.1 Dehydration Embrittlement

Dehydration embrittlement refers to brittle failure assisted by high fluid pore pressures that counteract the high normal stresses due to large overburden pressures. In the brittle failure of an intact material in a deviatoric stress field, failure occurs by growth and linkage of tensile (mode I) microcracks (Lockner et al., 1991; Scholz, 2002) oriented perpendicularly to the maximum compressive stress (**Figure 10**). The tensile cracks interact via the stress concentrations around their tips, self-organize, and coalesce into a through-going shear crack, which is oriented at an oblique angle to the microcracks and on which macroscopic failure and sliding takes place. A shear crack cannot grow in its own plane, but needs such mode I microcracks to create a process zone through which the shear crack can extend (Petit and Barquins, 1988). Formation

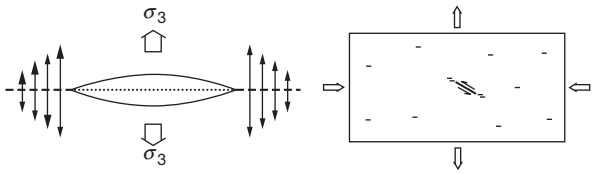


Figure 10 Relation between mode I microcracks and brittle shear failure. On the left, a tensile stress σ_3 opens a mode I microcrack perpendicular to σ_3. The microcrack concentrates tensile stress at its tips, promoting further growth. On the right, brittle shear failure occurs when such microcracks interact and link up. Macroscopic shear slip occurs at an oblique angle to the direction of greatest principal stress. Adapted from Frohlich C (2006) *Deep Earthquakes*. Cambridge, UK: Cambridge University Press; Green HW and Houston H (1995) The mechanics of deep earthquakes. *Annual Review of Earth and Planetary Sciences* 23: 169–213.

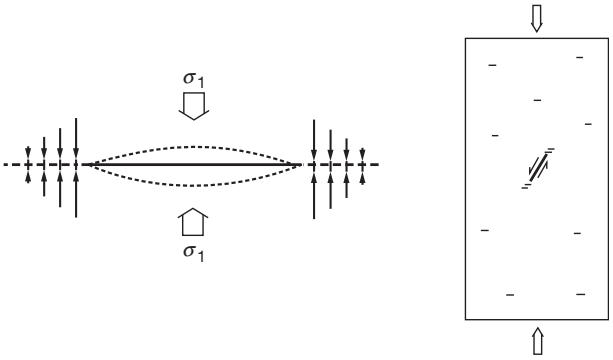


Figure 11 Relation between mode I anticracks and transformational faulting. On the left, a compressional stress σ_1 promotes transformation to a denser phase in a lens-shaped anticrack oriented perpendicular to σ_1. The anticrack concentrates compressional stress at its tips, promoting further growth. On the right, anticracks can interact, link up, and coalesce to form a plane on which macroscopic shear slip occurs at an oblique angle to the direction of greatest principal stress. Adapted from Frohlich C (2006) *Deep Earthquakes*. Cambridge, UK: Cambridge University Press; Green HW and Houston H (1995) The mechanics of deep earthquakes. *Annual Review of Earth and Planetary Sciences* 23: 169–213.

of mode I microcracks, and thus the resulting shear failure, is inhibited by the application of pressure. High fluid pressures (i.e., negative pressures) can hold open mode I microcracks that would otherwise be closed by overburden pressures. Deformation experiments in serpentinite showed the viability of this mechanism (Rayleigh and Paterson, 1965). In subducting slabs, the fluid may come from pore fluids squeezed from downgoing sediments or various metamorphic phase transitions in the subducting oceanic plate (Davies, 1999). Oceanic crust is pervasively hydrated at mid-ocean ridges, while oceanic lithosphere may be hydrated by fluid transport along outer-rise normal faults (e.g., Peacock, 2001). Bending-related faulting penetrates deep (at least 20 km) into the mantle at the Middle America trench and probably most subduction zones and likely promotes hydration of the deep crust and mantle surrounding the faults (Ranero et al., 2003).

The viability of dehydration embrittlement as a mechanism to generate deep earthquakes depends on the availability of fluids at relevant depths. The isotope Be^{10} is formed only at Earth's surface, so its presence in arc magmas is strong evidence that water in sediments is carried down by subducting slabs (Morris and Tera, 1989). There is abundant evidence that slabs may dewater via various phase transitions down to about 200–250 km depth (e.g., see Thompson, 1992). Laboratory experiments show that there are numerous possible mineralogies and transitions that could occur, many of them involving release of water (e.g., Table 1 of Omori et al., 2002). A popular candidate transition, due to the abundance of antigorite serpentinite, is the breakdown of serpentinite to olivine plus enstatite plus water. A special issue of the *Bulletin of the Earthquake Research Institute* (2001, vol. 76(3–4)) focuses on 'the role of water in earthquake generation.'

More recently, attention has turned to dense hydrous magnesium silicate phases (the so-called alphabet phases A through E), which could carry water deeper into the mantle as their stability fields range from ~200 km depth to within the lower mantle. Omori et al. (2004) analyzed *P–T* phase diagrams for hydrated peridotite and compared the depth distribution of dehydration reactions to that of intermediate and deep seismicity as a function of slab temperature. Based on similarities between the distribution of seismicity and dehydration reactions, they inferred that dehydration may act as a trigger for both intermediate and deep earthquakes. However, Barcheck et al. (2012) compared seismicity rate and calculated slab dehydration flux above 240 km (van Keken et al., 2011) but found little correlation between seismicity rate and dehydration flux (above 240 km) or seismicity rate and water remaining bound in slab minerals (in the 240–700 km depth range). The ability of the dense hydrous phases to generate deep earthquakes via dehydration embrittlement is not clear, because as slab material subducts following a low-temperature path to the higher pressures and temperatures at depth, the hydrous phases are capable of storing increasing amounts of water (e.g., Angel et al., 2001; Kohlstedt et al., 1996). Thus, net water would not be released during the phase transitions (although it might be available transiently), and the availability of free fluid to promote mode I cracks is questionable (Green and Marone, 2002).

4.13.3.2 Anticrack Faulting in a Metastable Phase

The possible relationship between deep earthquakes and the well-known phase transitions of olivine silicate to progressively denser forms through the upper mantle has long been pondered (e.g., see reviews by Frohlich, 2006a; Green and Houston, 1995; Kirby et al., 1996). However, seismic signals clearly show the deformation is shear, rather than implosive, in character (**Section 4.13.2.1**). Anticrack faulting provides a self-organizing mechanism analogous to fluid-assisted brittle faulting (i.e., dehydration embrittlement), by which the volume change and heat release of a phase transformation can trigger shear instability in a deviatoric stress field.

Anticracks are mechanically analogous to mode I cracks except that the signs of the displacement and stresses are reversed (Fletcher and Pollard, 1981). In the context of olivine phase transitions, anticracks comprise lenticular 'pods' of the transformed and denser phase (i.e., β- or γ-spinel). In contrast to mode I cracks, anticracks are oriented perpendicular rather than parallel to the maximum compressive stress (**Figure 11**).

Similarly to mode I cracks, the stresses induced around anticrack tips promote the lengthening or extension of the anticrack, that is, high normal stresses near their tips promote further phase transformation, which leads to their crack-like lenticular shape (**Figure 11**). Laboratory experiments on analog materials indicate that these structures can link up to form a through-going shear zone oriented not parallel, but obliquely, to the anticracks (**Figure 11**), on which macroscopic shear faulting takes place (e.g., Green and Burnley, 1989; Kirby et al., 1991; Schnubel et al., 2013; see Green and Houston, 1995 for a review). The shear instability requires that the starting phase exist metastably in the stability field of the final phase, so that anticracks filled with the final phase can grow slowly (e.g., aseismically). The final (equilibrium) phase was found in an extremely fine-grained form in the shear zone, as well as in the anticracks, suggesting that it lubricated the shear zone, permitting shear slip to occur. A key factor required for the instability or runaway is that latent heat be released by the transformation (Green and Zhou, 1996). As more anticracks form, heat is released, pushing the local region further away from the stability field of the metastable phase and resulting in a greater overstep of equilibrium conditions. Experiments on analog materials that have a positive volume change on transformation show that transformational faulting can proceed with either positive or negative volume change, but the reaction must be exothermic (Green and Marone, 2002; Green and Zhou, 1996; **Table 1**).

The viability of transformational faulting as a mechanism for deep earthquakes hinges on the presence of a sufficient amount of metastable phase (**Figure 12**). As the most abundant mineral in oceanic slabs is olivine, the most relevant phases for slabs in the upper mantle are α-olivine → β-spinel and α-olivine → γ-spinel. The equilibrium boundary of the olivine–spinel phase transition, normally at 410 km depth, is thermodynamically elevated in the vicinity of cold subducting slabs by 50–100 km as evidenced by an elevated 410 km seismic discontinuity near cold slabs (Vidale and Benz, 1992). Transformational faulting in olivine provides, therefore, a plausible mechanism only for earthquakes deeper than ~350 km. Thus, a key requirement for transformational faulting is the kinetic inhibition of transformation of a cold 'wedge' of the low-pressure phase, bounded approximately by a critical isotherm as the slab gradually subducts to deeper higher pressure–temperature environs. This requires sufficiently low temperatures inside the slab, discussed in **Sections 4.13.4.1** and **4.13.6.1**. Further, the extent to which transformation will continue to be inhibited on the long timescales relevant to subduction (e.g., several million years) is unclear.

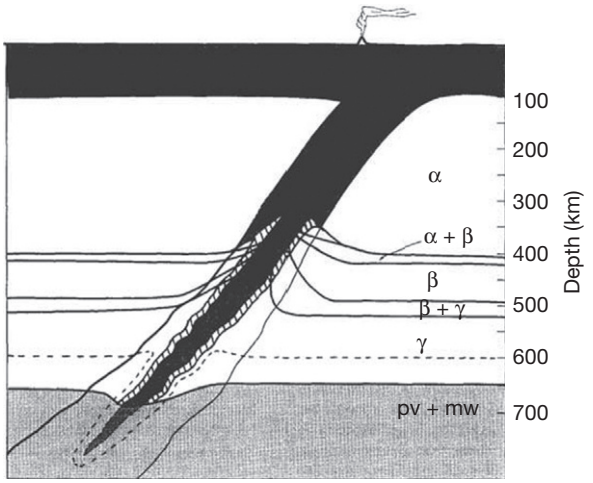


Figure 12 Schematic of metastable olivine wedge (dark region in core of slab below 350 km) in a cold subducting slab. Solid lines show the phase boundaries for the α, β, and γ polymorphs of mantle olivine. The hachured region represents a partially reacted region, proposed to be available for the generation of aftershocks under sufficiently high deviatoric stress (Green and Zhou, 1996). The dashed line shows the conditions for the metastable enstatite-to-ilmenite transformation. Adapted from Green HW and Zhou Y (1996) Transformation-induced faulting requires an exothermic reaction and explains the cessation of earthquakes at the base of the mantle transition zone. *Tectonophysics* 256: 39–56.

Vidale et al. (1991) and Koper and Wiens (2000) showed that the low seismic velocities associated with a metastable olivine wedge should distort the waveforms of some of the phases that travel through it. Subsequently, some evidence for low velocities consistent with a metastable wedge has been found in the waveforms and travel times of deep earthquakes. Iidaka and Suetsugu (1992) inferred a seismically slow core of the subducting Pacific slab beneath Japan from travel time analysis. Kaneshima et al. (2007) fit P waveforms and travel times of deep earthquakes in the Marianas slab with a 5% low-velocity anomaly. They interpret the anomaly as a metastable olivine wedge 25 km thick at 590 km depth; seismicity occurs outside and near one surface of the inferred wedge. Jiang et al. (2008) apply 3D ray tracing to fit travel times of deep earthquakes in the Japan slab with a 5% slow anomaly with a fingerlike shape, which they interpret to be a metastable olivine wedge. In this case most of the seismicity occurs within the inferred wedge. The high-density and high-quality seismic networks in Japan provided Kawakatsu and Yoshioka (2011) the opportunity to use receiver function analysis to directly image the subduction zone under Japan. The image features a 30 km thick dipping low-velocity layer in the same region deep under Japan that Iidaka and Suetsugu (1992) identified as slow. The elevated 410 km and depressed 660 km velocity discontinuities are clearly imaged as well, resembling those in **Figure 12**. **Section 4.13.5.1** outlines the evidence from deep double seismic zones in Tonga and Izu Bonin for metastable olivine wedges there.

In addition to the high-pressure polymorphs of olivine, phase transitions of other slab minerals could produce transformational faulting, although they are significantly less abundant than olivine. Hogrefe et al. (1994) reported that enstatite should persist metastably to greater depths than olivine and proposed that enstatite → ilmenite, an exothermic reaction with a large volume decrease (12%), could generate deep earthquakes via transformational faulting of metastable enstatite. Although enstatite comprises only ~25% of the slab by volume, that may result in a sufficient amount of metastable phase to permit transformational faulting.

Thus, this mechanism provides a natural explanation of the abrupt cessation of seismicity at the base of the upper mantle. Transformational faulting would not be expected to occur in the lower mantle because the possible reactions are

either endothermic (γ-olivine → perovskite + magnesiowustite and ilmenite → perovskite) or disproportionation (α-olivine → perovskite + magnesiowustite and garnet → perovskite phases) that require diffusion on the scale of the grain size and thus likely run too slowly to propagate a fault (Green and Zhou, 1996).

It is instructive to consider the similarities between dehydration embrittlement and transformational faulting (cf. **Figures 10** and **11**). In both cases, mode I tensile cracks or compressional anticracks filled with fluid or transformed material, respectively, self-organize in a deviatoric stress field via the stress concentrations at their tips to form a throughgoing shear zone (Green and Zhou, 1996). In both cases, shear zones are oriented, not in the plane of the mode I features, but about 30° to the direction of maximum compressive stress (**Figures 10** and **11**). Given this fundamental mechanical similarity despite large differences in the microphysics, together with the way that dynamic stresses drive rupture, the lack of dramatic changes in earthquake source properties with depth (e.g., stress drop, duration, and double-couple nature) is not necessarily indicative of the operation of only a single physical mechanism of failure.

4.13.3.3 Thermal Shear Instability

Although transformational faulting is, in a sense, a type of thermal shear instability (in that heat is released), in the deep earthquake literature, the term 'thermal shear instability' refers to shear localization produced by a positive feedback between temperature-dependent rheology and shear deformation that generates viscous heating. The nature of the feedback is such that under certain conditions, the localization of shear strain increases and accelerates exponentially, leading to apparently abrupt failure on a shear zone.

Variants of this mechanism have been proposed as a mechanism for deep earthquakes for many years (e.g., Griggs and Handin, 1960; Hobbs and Ord, 1988; Ogawa, 1987). Recently, Karato et al. (2001) observed that the most deformed slabs tend to be cold slabs. In their study of the dependence of slab rheology on temperature, Karato et al. (2001) incorporated the evolution in grain size associated with the transformation of olivine to spinel, plastic flow laws of the two phases, and a slab thermal model with latent heat release; three deformation mechanisms for olivine and spinel were included. The rheological structure of the slab is complex with weak, fine-grained spinel in the cold core surrounded by stronger, coarse-grained spinel regions on both sides. The cooler slabs are essentially weaker, and rapid deformation of a weak, cool slab focuses stress onto the strong regions. In this model, shear instability can occur in the strong regions if strain rate is large; thus, earthquakes could occur in the regions surrounding the weak fine-grained core. An instability parameter defines the regions subject to thermal shear instability (**Figure 13**). A 40–60 km thick zone is available to host deep earthquakes below ~450 km depth.

Kelemen and Hirth (2007) developed a numerical model of a viscous shear heating instability in a preexisting fine-grained shear zone, incorporating experimental olivine flow laws for

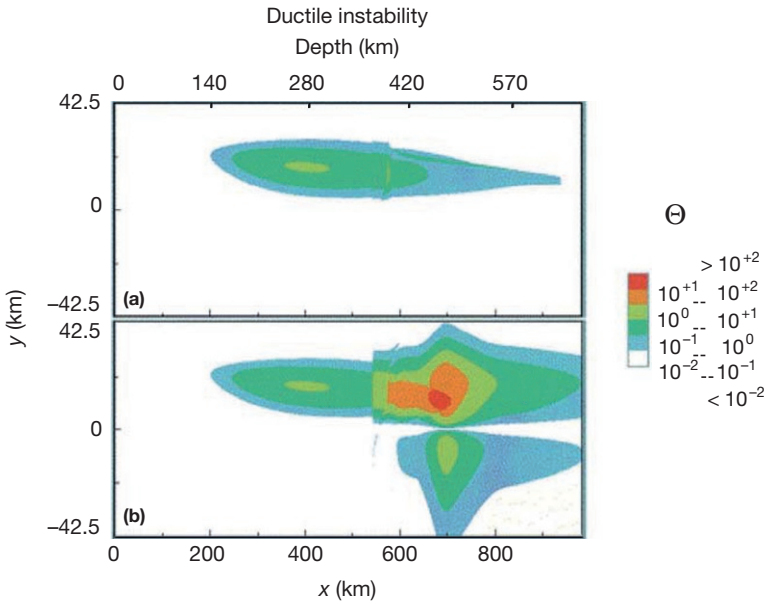


Figure 13 Portions of a slab subject to thermal shear instability. The assumed subduction rate is 10 cm year^{-1}. Distribution of the instability parameter Θ (a) when latent heat release and grain-size reduction are ignored and (b) when these effects are included (Karato et al., 2001). Thermal shear instability (and earthquake nucleation) is predicted to occur where the instability parameter exceeds 1. Note that the upper horizontal scale shows depth, while the lower gives distance along the dipping slab. The vertical scale shows distance perpendicular to slab dip. The thickness of the seismogenic regions can be up to 40–60 km. A cold slab is predicted to have a weak fine-grained spinel core; Θ exceeds the critical value for instability in the high-strength regions surrounding the core. Thus, a deep double seismic zone of limited extent is possible. Adapted from Karato S, Riedel MR, and Yuen DA (2001) Rheological structure and deformation of subducted slabs in the mantle transition zone: Implications for mantle circulation and deep earthquakes. *Physics of the Earth and Planetary Interiors* 127: 83–108.

dislocation creep, diffusion creep, and grain boundary sliding, as well as low-temperature plasticity. The model can produce repeated very rapid drops in strain in a single shear zone. The temperature range in which this periodic behavior operates (~600–800 °C) is similar to that in which most intermediate-depth earthquakes occur. However, the associated stress drops are ~1 GPa, far higher than those of intermediate-depth earthquakes.

If displacements are large and rapid enough, the fault surface or shear zone may melt. Fault slip on seismic timescales will typically generate heat much faster than it can be diffused away. If enough heat is generated by sufficiently large displacements across a sufficiently thin shear zone, temperatures may rise to the melting point. The temperature rise is proportional to the frictional stress and inversely proportional to the thickness of the shear zone (Kanamori et al., 1998). Estimates of temperature rise versus fault zone thickness for the M_w 8.2 1994 Bolivia earthquake suggest that melting would occur if the shear zone or slip zone was thinner than 0.3 m (Kanamori et al., 1998); it is difficult to assess whether such a thin shear zone was attained, but the shear instability process is inherently self-localizing. The relationship between rupture velocity, radiated seismic energy, moment, and static stress drop for the Bolivia event (in particular, the very low radiation efficiency, discussed in **Section 4.13.2.4.5**) implies that a large amount of energy was dissipated in the shear zone, perhaps by melting (Kanamori et al., 1998).

Shear localization in the shear instability mechanism is due to a progressive reduction in temperature-dependent viscosity due to heat generated by viscous shearing. If boundary conditions are suitable, for example, under conditions of constant stress, the viscosity reduction can lead to an increased rate of deformation that leads to greater reduction, and instability can result. But under typical conditions, strain may reduce the driving stress. Thus, it is not clear whether the actual boundary conditions in slabs (e.g., constant strain rate) can support a runaway positive feedback (Green and Marone, 2002). Thermal shear instability is known to occur in metals, but has not yet been definitively associated with earthquakes at any depth.

In the model of Karato et al. (2001), the observation of abrupt cessation of seismicity near 670 km in slabs subducting below that depth may be explained by superplasticity of perovskite in the lower mantle (Ito and Sato, 1991). But the observed low level of seismicity between ~250 and ~550 km is not well explained by the shear instability mechanism alone. In fact, the instability parameter of the mechanism predicts that seismicity should be strongest between 350 and 550 km depth (**Figure 13**), whereas observed seismicity increases below ~550 km (**Figure 1**). The instability parameter and the stress field must be considered together, but it is not clear that shear instability can easily explain the low levels of seismicity between 250 and 550 km, nor the resurgence of seismicity between 550 and 650 km.

4.13.4 Geophysical Setting of Subducting Slabs

4.13.4.1 Thermal State of Subducting Slabs

The thermal structure of slabs has been studied extensively, but only a few results are cited here. Thermal structure is central to the deep earthquake problem because all three potential mechanisms discussed in **Section 4.13.3** are controlled by slab temperature. Thermal control of the microphysics is exerted via dehydration rate, phase transformation rate, temperature-dependent rheology, or some combination of these. Syracuse et al. (2010) determined thermal structures down to 240 km depth for 56 subduction zone sections using numerical modeling of viscous flow with non-Newtonian, temperature- and stress-dependent rheology. Earlier, Kirby et al. (1996) related slab thermal structure to the depth of the deepest seismicity in the context of transformational faulting in metastable olivine. A more complete analysis of slab temperature by Devaux et al. (1997) included olivine–spinel reaction kinetics, latent heat release, and 2D heat conduction (up slab dip). Including these factors tends to increase slab temperatures and to reduce the length of a possible wedge of metastable olivine. The presence of metastable olivine at a given depth depends on the reaction kinetics and the minimum temperature in the slab at that depth. The critical temperature below which metastable olivine can be present may be weakly pressure-dependent (Kirby et al., 1991) and is usually estimated in the range 550–800 °C. Devaux et al. (1997) found that a metastable olivine wedge can exist down to ~660 km in only the coldest slabs (see their Table 3), but a few deep earthquakes occur as deep as ~660 km in apparently warmer slabs, such as South America. Mosenfelder et al. (2001) estimated somewhat higher temperatures and, thus, a shorter metastable wedge. However, the reaction kinetics are not well known. In sum, the viability of a metastable wedge to the base of the upper mantle is uncertain.

4.13.4.1.1 Thermal parameter

Since the minimum temperature at a given depth in a slab is of interest for most plausible physical mechanisms of deep earthquakes, the thermal parameter was developed as a crude way to compare this aspect of thermal structure for different subduction zones (Kirby et al., 1991). It is crude because it is based on a heat conduction model and does not include the effects of adiabatic heating or latent heats of transformation. Furthermore, the thermal parameter is less useful for studying the relation between intermediate-depth seismicity and dehydration. Much of the dehydration takes place at the top of the slab, and temperatures there are not described well by the thermal parameter.

The thermal parameter ϕ is defined as the product of a subducting slab's age at entering the trench, the convergence rate perpendicular to the trench, and the sin of slab dip; globally, it ranges from ~50 to ~12 000 in units of kilometer. ϕ is proportional to the inverse of minimum temperature in slabs at a given depth; thus, large ϕ signifies a relatively cool slab, and small ϕ, a warm one. Frohlich (2006b) recently developed an explicit relation of thermal parameter to minimum temperature, which does not include, however, the effects of adiabatic heating, latent heat, or 2D heat conduction.

Subduction zones with $\phi > \sim 3000$ are cool enough to host deep earthquakes (Kirby et al., 1991). This result was corroborated by Gorbatov and Kostoglodov (1997), who studied the relation of maximum depth of seismicity to ϕ for 110 locations along ten subduction zones. They find complex behavior (see their Figure 13), with a near-linear dependence for $\phi < 2500$

and depths of ~50 to ~300 km; then, the dependence of depth on ϕ weakens and then strengthens as maximum depth of seismicity increases rapidly from 300 to 600 km for just a small increase in ϕ. This sharp transition was also seen by Kirby et al. (1991), who argue that the reaction kinetics are such that a small decrease in slab temperature near $\phi \sim 3500$ results in a large change in the region of potential metastability. Finally, for ϕ higher than 6000 (i.e., the coolest slabs), the maximum depth of seismicity remains near 670 km for many subduction zones, raising the question why earthquakes do not continue into the lower mantle, which is addressed in **Section 4.13.3**. Gorbatov and Kostoglodov (1997) interpret the change in the ϕ-to-maximum-depth relation near 300 km as due to the olivine–spinel transformation, which has an equilibrium boundary near 350 km in subduction zones, in contrast to 410 km in ambient mantle. As latent heat of transformation is not included in ϕ, it may not represent temperatures well in that region, although if olivine remained metastable, ϕ would not change its relation to the minimum temperature.

In addition to the depth of the deepest seismicity, a number of earthquake properties seem related to ϕ. Wiens (2001) compared thermal parameter to magnitude–frequency relations (b-values) and aftershock occurrence rate (productivity) for events below 400 km. In cool slabs, the ratio of large to small events is smaller, and mainshocks produce more aftershocks, in contrast to the situation in warm slabs; the magnitude–frequency result is consistent with the results of Giardini (1988). Wiens (2001) also found a systematic relation of ϕ with the rupture properties of seven very large events ($M_w \geq 7.6$) deeper than 400 km, including source duration, static stress drop, rupture velocity, and maximum seismic efficiency. This approach was extended by Tibi et al. (2003a) who report source properties for 14 large earthquakes ($M_w \geq 7.0$) below 400 km. All of the previously mentioned properties, except stress drop, show a somewhat systematic relationship with ϕ. Large earthquakes in cooler slabs have shorter scaled durations, higher rupture velocities, and greater radiation efficiencies. The dependence indicates the sensitivity of the deep earthquake rupture process to slab temperature and was interpreted by Wiens (2001) and Tibi et al. (2003a) as support for the thermal shear instability mechanism. However, Persh and Houston (2004a) and Persh and Houston (2004b) compared ϕ to scaled duration and aftershock productivity for data sets of 111 and 290 earthquakes deeper than 100 km. The comparisons were made in three depth intervals (100–350 km, 350–550 km, and 550–700 km), but no clear relationship with ϕ was seen within the scatter. Nor was a clear relationship between rupture velocity and ϕ seen by Suzuki and Yagi (2011). Perhaps the largest events (those on which Wiens, 2001 and Tibi et al., 2003a are based) have a more systematic relationship to temperature. Also, for a given slab, values of ϕ obtained in different studies can differ by up to several thousand, mainly due to different estimates of slab age, rate, and dip (Frohlich (2006a), Figure 4.15). It would be useful to have modern, consistent, and comprehensive estimates of thermal parameter and, more importantly, slab thermal structure for all subduction zones, although uncertainty as to whether and where a metastable phase or, conversely, latent heat of transformation is present leads to significant uncertainty in temperature.

4.13.4.2 Stresses and Focal Mechanisms

The stress within a vertical column of material under its own weight provides an analogy to stresses in a subducting slab (Isacks and Molnar, 1971; Figure 5.4–11d of Stein and Wysession, 2003). In this simple thermomechanical model, tensional and/or compressional stresses are aligned along the dip of the slab. Down-dip compressional stresses dominate above the base of the upper mantle, if the slab is impeded by a higher-viscosity lower mantle; down-dip tensional stresses dominate at 100–200 km depth where the negatively buoyant slab essentially hangs from the surface.

Focal mechanisms from some subducting slabs are consistent with the thermomechanical model. Globally, mechanisms below 400 km are consistently down-dip compressional, while those above 300 km are often down-dip tensional, particularly in warmer slabs with no deep seismicity. Thus, this simple stress model has been invoked to explain the ubiquitous seismicity minimum around 250–450 km: as stress changes from down-dip tension just below the surface tectonic plates to down-dip compression above the base of the upper mantle, stress in the intervening region is expected to be low.

However, a number of slabs in the Western Pacific deviate from the predictions of the thermomechanical model and exhibit down-dip compression above the seismicity minimum (Chen et al., 2004; Green and Houston, 1995, see their Figure 7). Furthermore, in detail, focal mechanisms reveal locally heterogeneous strain, with different types of mechanisms coexisting in the same region (Giardini, 1992). Thermal expansion and equilibrium phase transitions in a compositionally layered slab (a 'petrologic' model) can lead to a heterogeneous stress state (Bina, 1996; Guest et al., 2003). Internal stresses due to phase transition-related volumetric changes are typically an order of magnitude greater than those due to buoyancy or thermal expansion (Devaux et al., 2000; Guest et al., 2003). The presence of a metastable wedge leads to additional spatial complexity (Devaux et al., 2000; Goto et al., 1987; Yoshioka et al., 1997). In their modeling of buoyancy-induced stress related to the olivine phase transitions, Yoshioka et al. (1997) found down-dip compression above 670 km, but below the metastable wedge, and maximum shear stresses up to 230 MPa along the wedge. Devaux et al. (2000) obtained differential stresses of ~1 GPa associated with a metastable wedge. Guest et al. (2003) obtained maximum shear stresses over 1 GPa due to equilibrium transitions with no metastability. Such high stresses are common in slab stress models; they are much higher than typical deep earthquake stress drops (1994 Bolivia excepted), which indicates that, in general, deep earthquake stress drops are not complete.

At intermediate depths, for the regions in which down-dip stresses dominate, Chen et al. (2004, see their Figure 5) find that down-dip compressional mechanisms are prevalent in old and fast slabs, whereas down-dip tension is prevalent in young and slow slabs (also see Figures 7 and 8 of **Chapter 4.18** by Ekstrom); that is, down-dip compression at intermediate depths tends to occur in the coolest and thus densest slabs, while the down-dip tension predicted by the previously mentioned thermomechanical model occurs in the less-dense slabs. This is inconsistent with the thermomechanical model but may

be consistent with a petrologic model. Chen et al. (2004) suggested that compressional stresses at intermediate depths may be generated in cold slabs by greater downwarping at the 670 km boundary, the potential presence of buoyant metastable olivine, or the inhibition of slab penetration at 670 km. However, it is not clear what is producing down-dip compression at intermediate depths in some subduction zones, as the results of Yoshioka et al. (1997) and Devaux et al. (2000), though complex, do not show this feature.

4.13.4.2.1 Fault-plane solutions and reactivated faults

A key aspect of the physical mechanism of deep earthquakes involves whether ruptures take place on new faults or on reactivated preexisting faults. A major issue for dehydration embrittlement involves how and to what depths slab mantle can become hydrated. One possibility involves fluid flow down the fault zones of large outer-rise normal faulting earthquakes prior to subduction. The fluids in such preexisting fault zones could be carried down with the subducting slab and released at depth to generate large intermediate or deep earthquakes.

Several recent studies have focused on whether faults formed by bending at the outer rise are reactivated at depth. In addition to paths for water to infiltrate deep into subducting slabs, large outer-rise faults could provide rheologically weaker zones. Seismic reflection imaging at the Middle America trench revealed pervasive bending-related normal faults, most of which dip trenchward (Ranero et al., 2003). Ranero et al. (2005) compared the strikes of normal faults at the outer rise to the strikes of intermediate-depth earthquakes in the Middle America and Chile subduction zones. They note that the strikes of the outer-rise faults can vary significantly along the trenches and the variation of the outer-rise fault strikes mimics that of the strikes of nodal planes of intermediate-depth earthquakes in the subducted slab below. The systematic similarity suggests that intermediate seismicity in these zones occurs on reactivated preexisting faults formed by plate bending at the trench.

At the Middle America and Tonga trenches (although not at Izu Bonin–Marianas), the vast majority of outer-rise faults dip toward rather than away from the trench (see references in Warren et al., 2007, 2008), even though either orientation would be consistent with the tensional stresses in the outer rise. Thus, accounting for the dip of the subducted slab, if these faults are reactivated at depth, then one would expect rupture mainly on near-vertical fault planes. The contrary, however, has been observed at intermediate depths by several different groups.

Some earlier inversions of body waveforms have found a tendency for intermediate-depth ruptures to propagate along strike and near-horizontally and thus prefer the near-horizontal nodal plane for the fault plane (e.g., Antolik et al., 1999; Tibi et al., 2002), although the resolution has been questionable. Warren et al. (2007) developed a technique to measure differential rupture durations from teleseismic P waveforms to obtain constraints on source directivity. This approach has allowed the fault plane to be identified in 30% of the ~190 events studied (Myhill and Warren, 2012; Warren et al., 2007, 2008). The vast majority of fault planes that can be reliably identified with this approach are near-horizontal and therefore inconsistent with the orientations of most of the outer-rise normal faults, which would yield near-vertical faults after subduction. Even in Izu Bonin–Marianas, where the outer-rise faults dip both toward and away from the trench, an explanation is needed to account for the lack of reactivation of trenchward-dipping faults (which would be near-vertical after subduction).

Warren (2010) assessed the resolution of the technique with synthetic tests, concluding that true rupture directions with both steep and shallow plunges are resolvable, given good station distributions and longer multipeaked time functions. Kiser et al. (2011) applied back-projection imaging to 22 intermediate-depth earthquakes and found a stronger tendency for ruptures to progress horizontally than vertically. Despite better resolution of shallowly dipping directivity than steeply dipping directivity due to a smaller range of angles between takeoff and rupture directions for the latter, the strong dominance of near-horizontal faulting found in the previous studies for intermediate quakes appears robust.

The predominance of near-horizontal rupture planes of intermediate earthquakes could have several interpretations. Perhaps faults are not reactivated but newly formed. In the Vanuatu subduction zone, stress inversions of intermediate-depth focal mechanisms indicate that new faults form if preexisting ones are not oriented favorably for failure (Christova and Scholz, 2003). The authors find significant differences between the distribution of nodal planes for intermediate-depth earthquakes and that for outer-rise events. The results imply that most intermediate-depth earthquakes in the Vanuatu slab lie on faults newly created at depth; the preexisting faults are only reactivated if they are oriented appropriately for the stress field. Myhill and Warren (2012) argued that the sum of down-dip stresses due to negative buoyancy and those due to unbending at depth in the Izu Bonin–Marianas subduction zone could favor slip on the horizontal conjugate fault plane over the vertical, just as bending in the outer rise of some trenches favors formation of trench-dipping faults. Warren et al. (2008) note that if ruptures are occurring on new faults, their near-horizontal orientations suggest the operation of an isobaric process. Kiser et al. (2011) proposed a mechanism involving interaction between pervasive postsubduction dehydration of the slab and the fabric imposed by preexisting faults formed at the outer rise. This explanation requires that some faults that dip away from the trench are generated at the outer rise; these are the ones more readily reactivated after subduction. Water migrating upward and along the slab could be preferentially trapped by near-horizontal preexisting faults that focus it deeper into the slab away from the slab surface, replenishing hydrous minerals along the fault. The fault could then slip by dehydration embrittlement. The conjugate near-vertical preexisting faults would be much less able to trap water and channel it downward toward the center of the slab against the vertical pressure gradient.

4.13.5 Seismic Structures in Slabs

4.13.5.1 Double Seismic Zones

In several subduction zones, intermediate-depth seismicity occurs in two layers up to 40 km apart, vertically separated by an aseismic region (**Figure 14**). Such double seismic zones have been detected at intermediate depths thus far at local sites in Japan, Tonga, Kamchatka, Alaska, Chile, New Britain,

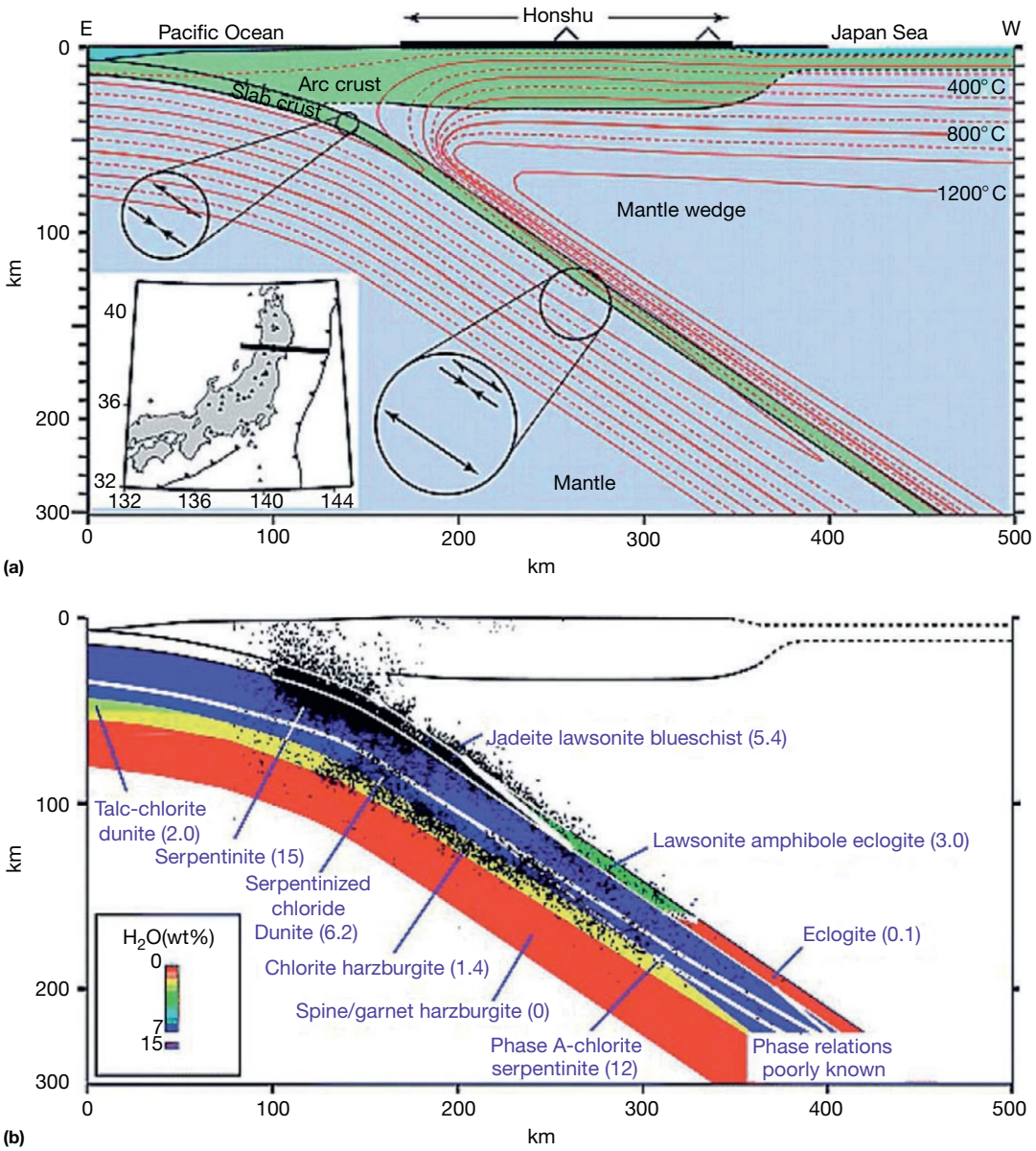


Figure 14 Cross sections depicting relations between thermal structure, predicted petrology, and earthquake locations in the Tohoku, Japan, subduction zone. (a) Thermal structure shown with solid isotherms every 200°. Offset circles show stress states inferred from Igarashi et al. (2001). (b) Comparison of petrology predicted by thermal structure with hypocentral locations. The lower zone of a well-developed double seismic zone follows the predicted locations of regions of serpentinite. Adapted from Hacker BR, Peacock SM, Abers GA, and Holloway SD (2003) Subduction factory 2. Are intermediate-depth earthquakes in subducting slabs linked to metamorphic dehydration reactions? *Journal of Geophysical Research* 108: 24627–24637. http://dx.doi.org/10.1029/2001JB001129.

New Zealand, Mexico, and Cascadia (e.g., Abers, 1992; Cassidy and Waldhauser, 2003; Comte and Suárez, 1994; Gorbatov et al., 1994; Hasegawa et al., 1994; Kawakatsu, 1986; Pardo and Suarez, 1995; Rietbrock and Waldhauser, 2004; Robinson, 1986). In the best-resolved double zones in northeastern Japan (Hasegawa et al., 1994) and Kamchatka (Gorbatov et al., 1994; Kao and Chen, 1994), the lower zone is seismically active between 30 and 180 km depth, below which it merges with the upper zone. Systematic analysis of a global seismicity catalog revealed evidence for bimodal distributions of seismicity relative to the slab surface consistent with double seismic zones in 30 segments of 16 subduction zones (Brudzinski et al., 2007). The separation ranges from 8 to 30 km with a tendency to increase with slab age; this result is consistent with the dehydration of basalt and antigorite enabling dehydration embrittlement in the upper and lower zones, respectively (Brudzinski et al., 2007).

Focal mechanisms indicate that the majority of seismic double zones are consistent with down-dip compression in the upper zone and down-dip tension in the lower zone. This is consistent with a model of stresses produced by unbending of the slab as it straightens out below the overriding lithosphere, and this model has provided the conventional explanation of the stresses causing double zones (e.g., Isacks and

Barazangi, 1977; Kawakatsu, 1986). But double-zone seismicity continues well after the slab has straightened out. Perhaps minor sagging could be involved. Furthermore, the double zones beneath Alaska (Abers, 1992) and New Zealand (Robinson, 1986) have both layers in down-dip tension, while for those beneath Mexico (Pardo and Suarez, 1995) and northern Chile (Comte and Suárez, 1994), the upper zone is in down-dip tension and the lower zone appears in down-dip compression. In northeastern Japan, Igarashi et al. (2001) detected a third seismic plane 5 to 10 km above the down-dip compressional upper plane; the third and uppermost plane appears to lie at and just beneath the top surface of the descending slab. It consists of interplate low-angle thrust earthquakes near the top surface of the descending slab from 30 to 60 km depth, transitioning to intraplate normal faulting (down-dip tensional) mechanisms within the upper crust from 60 to 110 km depth (Igarashi et al., 2001; Kita et al., 2006). A similar three-plane zone may be present in Kamchatka, where Gorbatov et al. (1994) found a handful of events analogous to the down-dip tensional events in the uppermost (third) plane. Thus, there must be extensional stresses around the slab surface, which cannot be explained by large-scale plate unbending alone. Stress heterogeneities due to petrologic layering of the slab may provide an explanation, that is, the extensional stress state of the uppermost (third) plane could be associated with the formation of eclogites, while the stress state of the other planes may be largely caused by unbending (Kita et al., 2006; Wang, 2002). Kita et al. (2010) inverted slab focal mechanisms in northeastern Japan for stress orientations to estimate the location of the neutral plane between down-dip compressional and down-dip extensional layers. They obtained a shallower neutral plane below the slab surface in Hokkaido than in Tohoku and suggest that a smaller, less-developed metastable olivine wedge under Hokkaido provides less buoyancy to counteract the general down-dip extension at intermediate depths exerted by slab pull. Kao and Chen (1994) proposed that the compression responsible for the lower plane of the Kamchatka double zone is transmitted to intermediate depths from the base of the upper mantle.

For many double zones, it is clear that the lower seismic zone must lie deep within the subducting mantle lithosphere, raising the question of how originally dry oceanic mantle lithosphere can become hydrated. Peacock (2001) suggested that normal faulting associated with bending at the outer rise and trench allows infiltration of water tens of kilometers into the oceanic lithosphere. Evidence for pervasive bending-related normal faulting that penetrates at least 20 km deep into the plate is seen at sites along the Middle America trench (Ranero et al., 2003). Near the trench, reflection data reveal about 1.5 faults per kilometer with offsets of 100–1000 m. An estimate of the amount of water in the slab subducting at the Middle America trench suggests that the serpentinized mantle portion of the slab may hold as much chemically bound water as the slab's crustal portion (Ranero et al., 2003). Furthermore, the patterns of orientation of such faulting are similar to those of intermediate-depth seismicity in the Middle America and Chile subduction zones (Ranero et al., 2005), suggesting that intermediate seismicity occurs on preexisting faults.

Hacker et al. (2003) compared the morphology of intermediate-depth seismicity for four subduction zones (Cascadia, Nankai, Costa Rica, and Tohoku) with the locations and water content of hydrous minerals, predicted based on slab age and thermal structure. The depth and geometry of the seismicity are consistent with the predicted distribution of several hydrous phases. In the coldest zone, Tohoku, double planes of seismicity surround the cold core of the slab: the zones approach one another, nearly following isotherms, but actually cutting across them at a shallow angle (see also Kita et al., 2006). Antigorite serpentinite, one of the most stable hydrous phases, is the main component in cooler slabs at intermediate depths and is believed to dehydrate progressively down to about 200 km. Yamasaki and Seno (2003) obtained a similar result analyzing double seismic zones in six subduction zones. These studies provide strong evidence for dehydration embrittlement as the mechanism for intermediate earthquakes down to perhaps 250 km (Hacker et al., 2003 and Yamasaki and Seno, 2003). Double seismic zones appear easier to detect in cooler slabs, probably because the two layers are farther apart.

Double seismic zones have also been observed below 300 km depth. Wiens et al. (1993) relocated hypocenters with P, pP, and PKP arrivals and resolved a deep double seismic zone in Tonga from 350 to 460 km, with about 30 km separating the two layers. The stress state for the Tonga double zone is roughly opposite to the state typical at intermediate depths. Iidaka and Furukawa (1994) detected two planes about 25 km apart in the Izu Bonin subduction zone using S–P converted phases at the upper boundary of the slab. The double zone extends from about 300 to 450 km depth, that is, below the seismicity minimum, which in the Izu Bonin slab occurs in the anomalously shallow range of 200–300 km. The upper layer lies 15–20 km below the top surface of the slab. The double zone occurs in the same region that Iidaka and Suetsugu (1992) inferred a seismically slow core of the slab from travel times (see discussion in Green and Houston, 1995). Both these cases were interpreted as evidence for transformational faulting on the upper and lower edges of a metastable olivine wedge (Iidaka and Furukawa, 1994; Wiens et al., 1993). A comparison of the modeled stresses due to a metastable wedge to the characteristics of the deep double zone in Tonga suggested that the lower seismic zone may reside well below a hypothesized metastable wedge, which would imply the operation of two mechanisms of rupture (Guest et al., 2004).

4.13.5.2 Planar Shear Zones

Planar seismicity structures have been detected below 400 km in several slabs. In local sites in the Tonga, Japan, Bonin, Marianas, Java, Banda, and South American slabs, seismicity is concentrated along large-scale planes, which align with a focal mechanism nodal plane, suggesting that large throughgoing shear zones host the seismicity (Giardini and Woodhouse, 1984; Lundgren and Giardini, 1992). These features are best developed in Tonga where they can extend over 200 km and cut through the seismically active portion of the slab (Giardini and Woodhouse, 1984). In some cases, active conjugate shear zones intersect at $\sim 90^\circ$ (Lundgren and Giardini, 1992). The sense of deformation is related to slab contortion, buckling, and imbrication in Tonga (Giardini and Woodhouse, 1984) and in general reflects resistance to slab penetration at 670 km and flattening of the slab to a more horizontal position (Lundgren and Giardini, 1992). Tibi et al.'s (2001) relocations of 17 aftershocks of the 1995 M 7

Mariana earthquake (586 km depth) span a plane about 20 km, across the entire thickness of the seismically active slab.

4.13.5.2.1 Repeating earthquakes

Recent more precise earthquake locations reveal structures similar to planar shear zones on a smaller scale. Wiens and Snider (2001) detected clusters of M 4–5 earthquakes with identical waveforms, 560–600 km deep in the Tonga slab. Precise relocations using waveform cross correlation yield location accuracies of 1–3 km and reveal that the events locate on shear zones 10–30 km in length and up to only 2–3 km thick. Three events are collocated within 1 km and their rupture zones appear to overlap, so they could represent repeated failure of same fault surface. The authors interpret the results in terms of runaway thermal shear instabilities and suggest that earthquake doublets within a few hours represent occurrence of the second event before heat generated by first event can diffuse away (Wiens and Snider, 2001). However, location accuracies of 1–3 km, while a large improvement, are still insufficient to rule out transformational faulting, as the apparently collocated events could occur on parallel en echelon faults that cannot be spatially distinguished.

4.13.5.3 Outboard Earthquakes and Recumbent Slabs

In several locations, a few deep earthquakes have been observed well outside the main Wadati–Benioff zone seismicity; these events, which occur farther from the trench than the dipping Wadati–Benioff zone, are termed 'outboard earthquakes.' They appear to indicate flattening or horizontal deflection of some deep slabs near the base of the lower mantle.

Lundgren and Giardini (1994) documented several outboard earthquakes at about 600 km depth in the Izu Bonin, Kurile, and Chile subduction zones, which lie 150–300 km away from the main Wadati–Benioff zone. In the first two zones, Okal (2001) established the continuity of the subducting lithosphere between the main seismic zone and the outboard event, consistent with horizontal deflection of the actively subducting slab. Lundgren and Giardini (1994) also described more extensive subhorizontal distributions of seismicity in Banda and Tonga–Fiji, which appear to delineate nearly flat-lying slab remnants recumbent near the base of the transition zone in these regions. Seismicity in the latter region was scrutinized by Okal and Kirby (1998), who reached a similar conclusion. Deep seismicity near Tonga–Fiji has been recognized as highly spatially complex for decades and has been interpreted as indicating a very contorted slab. However, Chen and Brudzinski (2001) thoroughly analyzed the geometry and focal mechanisms of seismicity in the Tonga–Fiji region and resolved that two remnants of previously subducted lithosphere are detached from the currently subducting slab; portions of the remnants are nearly horizontal and up to 1000 km away from the actively subducting slab, while a portion lies above the actively subducting slab (**Figure 15**). The detached remnants may be associated with an episode of subduction along the fossil Vitiaz trench 5–8 Ma. Stress orientations from focal mechanisms of outboard earthquakes and the remnant slab seismicity are distinct from orientations in the main Wadati–Benioff zone and are not those expected if the detached remnants were a stress guide. Their P-axes lie closer to the slab normal, whereas deep in a Wadati–Benioff zone seismicity is typically down-dip compressional.

Thus, rather than sinking undeformed into the lower mantle, in some subduction zones, lithosphere is deflected, retained in the transition zone, and, most saliently, still capable of earthquakes. There must, therefore, exist a localized source of stress in the subhorizontal portions, rather than

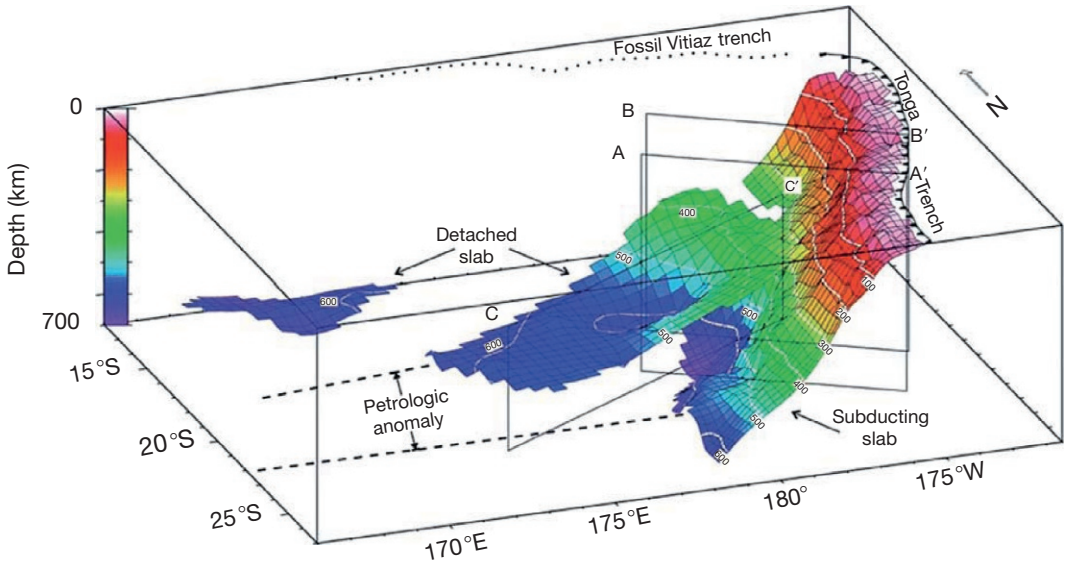


Figure 15 A three-dimensional view of the Tonga slab for the region between 15 °S and 25 °S. Slab contortions and detached remnants are apparent. Three surfaces were constructed from the hypocentral locations, representing the actively subducting lithosphere, a detached remnant of slab, and an isolated cluster of earthquakes northeast of the Vanuatu trench. Dashed lines indicate the region of an inferred petrologic anomaly, in which anomalously slow wave speeds are detected. White contour lines (100 km in depth) are shown. Adapted from Chen W-P and Brudzinski MR (2001) Evidence for a large-scale remnant of subducted lithosphere beneath Fiji. *Science* 292: 2475–2479.

slab pull or resistance to penetration (Brudzinski and Chen, 2005). Based on their global analysis of patterns of strain in subhorizontal slabs at both intermediate depths and deeper, Brudzinski and Chen (2005) emphasized the need for locally variable sources of stress to generate seismicity in subhorizontal portions of slabs.

Significantly, the subhorizontal swaths of seismicity that delineate slab remnants in the Tonga–Fiji region are not associated with a fast velocity anomaly such as is typical for rapidly subducting zones, although adjacent regions without subhorizontal seismicity exhibit 2–3% faster P and S velocities (Brudzinski and Chen, 2003; Chen and Brudzinski, 2001). The line of demarcation is sharp (Brudzinski and Chen, 2003). The interpretation is that the relatively cold temperature of the slab remnants must have been counteracted by a petrologic anomaly, a compositional or mineralogical difference, such as the presence of metastable olivine. Indeed, the preferred explanation of Brudzinski and Chen (2003) is that metastable olivine is retained in the remnant slabs and generates their seismicity. This explanation is consistent with a number of observations (e.g., Brudzinski and Chen, 2003, 2005). However, it contradicts available thermal constraints on olivine metastability (Devaux et al., 1997 and Mosenfelder et al., 2001), because after 5–8 My in the transition zone, the remnant slabs must be relatively warm compared to the adjacent currently subducting lithosphere. Whether remnant slabs could contain metastable olivine remains an open question.

4.13.6 Implications for Seismogenesis

4.13.6.1 Viability of the Three Main Physical Mechanisms

Here, the author outlines key observations and inferences consistent and inconsistent with the proposed mechanisms of deep earthquake rupture discussed in **Section 4.13.3**.

Dehydration embrittlement provides a good explanation of the frequency and location of intermediate-depth seismicity from ~50 down to ~250 km. The exponential decrease with depth in earthquake frequency from 50 to 350 km (**Figure 1**) may be indicative of the gradual exhaustion of water as metamorphic reactions proceed. The morphology of intermediate-depth double seismic zones is consistent with the dewatering of hydrous phases predicted from subduction zone thermal structures (Hacker et al., 2003; Yamasaki and Seno, 2003). Antigorite serpentinite is the most common constituent of subducting lithosphere at intermediate depth in cold slabs and its behavior is well known (Rayleigh and Paterson, 1965). Omori et al. (2004) and van Keken et al. (2011) have discussed dehydration reactions and bound water in subducting slabs. Although dehydration reactions have been proposed as triggers for deep, as well as intermediate, earthquakes (Omori et al., 2004), Green and Marone (2002) inferred that dehydration embrittlement appears less likely below 250 km, because hydrous phases, if present, probably retain rather than release water as they transform.

A plausible mechanism has been put forward for pervasive hydration along bend-related faults at the outer rise (Ranero et al., 2003, 2005); perhaps the hydration is sufficiently pervasive in some cases to allow formation of new faults with different orientations (Christova and Scholz, 2003). Slab unbending does not provide an entirely satisfactory explanation for the stresses driving intermediate-depth seismicity, due to the variety and complexity of stress states observed in double zones. Indeed, Brudzinski and Chen (2005) suggested that patterns of strain in subhorizontal slabs at intermediate depth and deeper indicate the need for locally variable sources of stress.

The micromechanics of transformational faulting is analogous to that in dehydration embrittlement. Transformational faulting has been well demonstrated in the laboratory in analog materials (e.g., Green and Burnley, 1989; Kirby et al., 1991; and recently Schubnel et al., 2013), though less well in natural silicates. In concert with dehydration embrittlement operating above ~350 km, transformational faulting in metastable olivine below 350 km would explain seismicity's bimodal depth distribution and cessation at 670 km (Green and Zhou, 1996),and has been a conventional and popular explanation (Green and Houston, 1995). However, recent thermal modeling suggests that only the coldest slabs can host metastable olivine down to the base of the upper mantle (Devaux et al., 1997; Mosenfelder et al., 2001), although warmer slabs such as those in South America exhibit seismicity nearly as deep as those in Tonga. Unfortunately, the relevant reaction kinetics are not well known. The possibility of transformational faulting in other metastable phases, such as enstatite, should not be overlooked (Hogrefe et al., 1994). Metastable olivine provides an explanation for the seismic properties of an apparent remnant slab in the upper mantle beneath Tonga (Chen and Brudzinski, 2001). Most saliently, it has apparently been detected deep in the Pacific slab under Japan by high-resolution receiver function imaging (Kawakatsu and Yoshioka, 2011).

One of the main arguments advanced against transformational faulting in metastable olivine is that several recent very large and deep earthquakes exhibit fault planes that appear too large to fit into the expected narrow wedge of cold metastable material (e.g., McGuire et al., 1997; Silver et al., 1995). Furthermore, aftershocks nucleate in a wide zone outside the region of background seismicity (McGuire et al., 1997). This argument, although strong, is not completely definitive in view of the difficulty resolving the extent of rupture (e.g., compare geometry of rupture planes of the 1994 Bolivia earthquake determined by Chen, 1995; Ihmle, 1998; Kikuchi and Kanamori, 1994) as well as the possibility that slabs may thicken near 670 km due to resistance at the base of the mantle (Kirby et al., 1996).

The thermal shear instability mechanism may be consistent with the observation of large, very deep fault planes, although for warmer slabs such as those in South America, a large bending moment must be applied (Karato et al., 2001). However, the shear instability mechanism by itself does not explain the depth distribution of seismicity well, particularly for intermediate-depth earthquakes. Modeling (Karato et al., 2001) indicates that the shear instability mechanism may begin to operate at about 200 km depth. The instability parameter grows with depth and appears strongest between 350 and 550 km depth (**Figure 13**). In contrast, actual seismicity decreases exponentially from 50 to 350 km, is minimal from 350 to 500 km, and then resurges between 500 and 650 km (**Figure 1**). Deviatoric stresses in the slab must be considered in

conjunction with the instability parameter; they are sufficiently complex even for the equilibrium case (e.g., Guest et al., 2003; Yoshioka et al., 1997) that it may be possible that they could produce the resurgence of seismicity below 500 km. On the other hand, the model of Kelemen and Hirth (2007) operates only in the depth range of intermediate-depth earthquakes, but produces stress drops an order of magnitude too high, and does not explain the tendency for horizontal rupture in intermediate quakes nor their exponentially decreasing frequency with depth.

It is not clear what would stop an earthquake generated by shear instability. Faulting may be expected to run until driving stress falls below the viscous resistance, which must have plummeted to allow the instability. Static stress drops of most deep earthquakes are well below modeled deviatoric stresses in most cases, with 1994 Bolivia providing the main exception. Finally, the boundary conditions must be appropriate to accelerate shear and drive the instability.

Any of the earlier mechanisms might ultimately result in melting on the fault plane if displacement is large enough, but a progression to melting is most natural via thermal shear instability. Kanamori et al. (1998) argued that melting occurred in the 1994 Bolivia earthquake based on the large amount of energy apparently dissipated in the fault zone, as evidenced by the energy budget and the low rupture velocity. On the other hand, it seems odd that the Bolivia earthquake's radiated energy-to-moment ratio and scaled duration are similar to those of other deep and shallow earthquakes. The event's energy budget differs from that of other earthquakes primarily due to its high static stress drop, which is closely related to its low rupture velocity.

4.13.6.2 Notable Points and Speculations

It is clear that deep earthquakes are primarily double-couple (shear) sources and trigger fewer aftershocks than shallow earthquakes. Variations in some rupture processes with depth are detectable in the scattered data. Aftershock productivity, scaled duration, and source complexity vary with depth. Although many rupture properties of deep earthquakes are not distinctly different from shallow quakes, it is likely that rupture velocity varies more for deep than for shallow earthquakes. For shallow faulting, rupture velocity usually falls between 0.6 and 0.9 of the shear-wave velocity, whereas for deep earthquakes, it varies more widely from 0.2 to near 1.0 of the shear-wave velocity.

What stops earthquake rupture is a long-standing question for deep and shallow events alike. The end of rupture is hard to resolve in seismic waves due to noise, but stacking has increased resolution. In some cases, deep earthquake rupture seems to stop abruptly (see Figure 1b of Vidale and Houston, 1993). It is not obvious how observations of abrupt termination can be reconciled with mechanisms like shear instability and melting, which appear to call for a gradual decay of moment-release rate and a complete stress drop. Stress drops, while intrinsically variable and difficult to determine, appear to be much lower than many estimates of deviatoric stresses at depth in subducted lithosphere.

A comparison of very deep South American seismicity with that in Tonga suggests that (below 400 km) conditions promoting the nucleation of deep rupture differ from those promoting its propagation. The subduction zone with the most earthquakes, Tonga, is not the site of the largest ruptures, which are hosted in the relatively warm South American subduction zone. This contrast is reflected in the different relation between frequency and magnitude (i.e., *b*-value) for these two zones, with other subduction zones that possess deep seismicity falling intermediate between the two end-members (Giardini, 1988; Wiens, 2001). Notably, that contrast does not occur for intermediate-depth earthquakes. Thus, in a way not seen for intermediate-depth events, the propensity to nucleate or propagate rupture of events below 400 km appears controlled by the thermal setting, as are some rupture properties (Tibi et al., 2003a; Wiens, 2001).

The largest deep events occur in isolated seismically quiet regions with anomalously few aftershocks. They have low rupture velocities, a relatively weak initial phase, and a tendency to rupture sideways (Estabrook, 1999). Their isolation leads to a prediction that over the decades, more very deep giant earthquakes are likely to strike in nearby currently quiet zones, including regions north of the 1970 Colombia earthquake. These observations raise the possibility that different mechanisms may operate to generate deep earthquakes in different zones (Wiens and McGuire, 1995) or different mechanisms, with different dependencies on pressure–temperature conditions, may be important during different phases of rupture (e.g., nucleation vs. propagation).

More generally, the possibility that multiple mechanisms generate deep earthquakes must be considered. Which of the mechanisms is operative could depend on depth, subduction zone (e.g., slab age, convergence rate, and dip), or earthquake size. The author has already set forth the conventional model that dehydration embrittlement is responsible for intermediate-depth events and transformational faulting in metastable olivine for events below 350 km (Green and Houston, 1995; Kirby et al., 1996). Persh and Houston (2004b) interpreted the dependency on depth seen in some source properties as suggestive of a change in mechanism near 550 km, in addition to the change around 350 km. In the region between ~350 and 550 km, scaled duration, complexity, and aftershock productivity are anomalously long, high, and low, respectively, and the changes at ~550 km appear abrupt (Persh and Houston, 2004a,b). Furthermore, Suzuki and Yagi's (2011) finding of relatively high rupture velocities ($>0.5\beta$) only between 530 and 610 km depth is also suggestive of a change in the mechanism of rupture near the depth of ~550 km. Possible mechanisms below ~550 km include shear instability, melting, and transformational faulting in metastable enstatite. This scenario circumvents the apparent difficulty preserving metastable olivine below 550 km for most slabs (e.g., Devaux et al., 1997), as well as the difficulty connected with rupture planes larger than the expected thickness of metastable olivine (e.g., McGuire et al., 1997; Silver et al., 1995). On the other hand, the contrast in behavior of the deepest seismicity between Tonga and South America suggests the possibility that transformational faulting in metastable olivine operates in Tonga below 350 km, while a different mechanism operates deep in the warmer South American slab – for example, shear instability or transformational faulting in metastable enstatite. The complete lack of seismicity

between about 325 and 500 km depth along the length of the South American slab is striking compared to other zones. Finally, it is also plausible that in very large, deep earthquakes, rupture could initiate by one mechanism but progress to another (e.g., shear instability or melting) after the initial displacement generated sufficient heat. This would result in a variation in the mix of physical mechanisms for different-sized earthquakes. The energy budget and waveform shape suggest that this may have occurred for the 1994 Bolivia earthquake and, if so, probably for the other South American very large deep earthquakes (Estabrook, 1999; Kanamori et al., 1998).

Other combinations are possible, for example, in Tonga, shear instability and, in South America, transformational faulting in metastable enstatite, followed by shear instability and melting for the largest events. The combinations mentioned here were chosen so that rupture initiation would occur more readily in Tonga and propagation more readily in South America, due to the effects of different thermal structures or concentrations of a metastable phase; that is, transformational faulting might initiate less readily in metastable enstatite than in metastable olivine, because the enstatite–ilmenite component comprises only ~25% of the slab mineralogy, whereas the olivine–spinel component comprises the bulk of the slab mineralogy. Similarly, warmer temperatures favor enhanced propagation via shear instability or melting (once slip is initiated).

At this time, no single physical mechanism appears capable of generating all seismicity below 100 km, or even below 200 km, and explaining observations of earthquake depth distribution, stress orientations, *b*-values, aftershock productivity, and rupture processes. In particular, dehydration embrittlement cannot readily explain both the contrast between subduction zones in magnitude–frequency relations for events below 350 km discussed earlier and the lack of contrast for events above 350 km, nor both the depth distribution and stress state. Given the subduction zone complexity that is emerging from recent studies, perhaps diversity in the mechanisms of rupture generation should be expected; perhaps it is also suggested by the variation in rupture velocities. Subducting slabs are inherently complex in mineralogy, rheology, and stress due to their petrologic layering. The contribution of minor mineral components to the complex rheology and stress may be significant, and the mineralogy of subducting lithosphere and crust appears to be poorly known in detail. Rheology may be controlled by factors beyond thermal structure or the minimum temperature at a given depth; that is, an old, but slow, slab may have a different rheology than a young, but fast, slab, even if they are predicted to have the same minimum temperature at a given depth. Furthermore, the varied deformation history of subducted slabs, for example, the degree of bending at the trench or shear in the mantle, may affect their rheology.

Preexisting faults likely play a very important role in intermediate-depth seismicity (Ranero et al., 2003, 2005), but if they are not oriented sufficiently favorably, new faults can form (Christova and Scholz, 2003). Below 400 km, seismicity organized in planar structures appears to delineate fault zones or their physical equivalents at 400–700 km depth (e.g., Giardini and Woodhouse, 1984; Lundgren and Giardini, 1992). The occurrence of deep shear zones that persist as stable features of a deep subducted slab suggests that preexisting faults play a role in deep seismicity, although they need not have formed at Earth's surface. These structures are not predicted by transformational faulting, and depending on their geometry, they could be inconsistent with that mechanism. Their existence is evidence of a type of heterogeneity not often considered in modeling the rheology and stress of deep subduction zones.

4.13.7 Concluding Thoughts

Although evidence has accumulated over the past decade that intermediate-depth earthquake is likely generated by dehydration embrittlement, in my view, deeper earthquakes will not give up their secrets easily. Considering that the physical processes that occur during shallow rupture are still not fully understood, it is not entirely surprising that the more remote phenomena remain ambiguous. Progress has been slowed by the variability of earthquake source properties, by the poorly known mineral physics due to the difficulty of achieving in the laboratory conditions at great depth, and perhaps most of all by the lack of robust links between earthquakes source properties and physical mechanisms. Models connecting properties of the source to possible mechanisms lack a firm basis and are not yet well developed. Integration of seismological, laboratory, and modeling efforts is needed to bridge the stubborn gap between source properties, which are extracted under strong assumptions and possess substantial intrinsic variability, and physical mechanisms of rupture generation, which are as yet neither well understood nor constrained. Heterogeneity in slab structure, stress, and rheology must be considered. Comparative studies of many events in many subduction zones will help overcome scatter and clarify systematic relationships. It would be helpful to have a base of systematic up-to-date fundamental information, such as thermal structures (perhaps versions both with and without olivine metastability) using known tectonic parameters for all subducting slabs. More precise earthquake locations and focal mechanisms will be valuable. An organized effort to provide these on a global basis would probably drive significant progress.

References

Abers GA (1992) Relationship between shallow- and intermediate-depth seismicity in the eastern Aleutian subduction zone. *Geophysical Research Letters* 19: 2019–2022.

Angel RJ, Frost DJ, Ross NL, and Hemley R (2001) Stabilities and equations of state of dense hydrous magnesium silicates. *Physics of the Earth and Planetary Interiors* 127: 181–196.

Antolik M, Dreger D, and Romanowicz B (1999) Rupture processes of large deep-focus earthquakes from inversion of moment–rate functions. *Journal of Geophysical Research* 104: 863–894.

Barcheck CG, Wiens DA, van Keken PE, and Hacker BR (2012) The relationship of intermediate- and deep-focus seismicity to the hydration and dehydration of subducting slabs. *Earth and Planetary Science Letters* 349: 153–160. http://dx.doi.org/10.1016/j.epsl.2012.06.055.

Beck SL, Silver P, Wallace TC, and James D (1995) Directivity analysis of the deep Bolivia earthquake of June 9, 1994. *Geophysical Research Letters* 22: 2257–2260.

Bina CR (1996) Phase transition buoyancy contributions to stresses in subducting lithosphere. *Geophysical Research Letters* 23: 3563–3566.

Blanco MJ and Spakman W (1993) The P-wave velocity structure of the mantle below the Iberian Peninsula: Evidence for a subducted lithosphere below southern Spain. *Tectonophysics* 221: 13–34.

Brace WF and Kohlstedt DL (1980) Limits on lithospheric stress imposed by laboratory experiments. *Journal of Geophysical Research* 85: 6248–6252.

Brudzinski MR and Chen W-P (2003) A petrologic anomaly accompanying outboard earthquakes beneath Fiji–Tonga: Corresponding evidence from broadband P and S waveforms. *Journal of Geophysical Research* 108: 2299. http://dx.doi.org/10.1029/2002JB002012.

Brudzinski MR and Chen W-P (2005) Earthquakes and strain in subhorizontal slabs. *Journal of Geophysical Research* 110: B08303. http://dx.doi.org/10.1029/2004JB003470.

Brudzinski MR, Thurber CH, Hacker BR, and Engdahl R (2007) Global prevalence of double Benioff zones. *Science* 316: 1472–1474. http://dx.doi.org/10.1126/science.1139204.

Brustle W and Muller G (1987) Stopping phases in seismograms and the spatiotemporal extent of earthquakes. *Bulletin of the Seismological Society of America* 77: 47–68.

Cassidy JF and Waldhauser F (2003) Evidence for both crustal and mantle earthquakes in the subducting Juan de Fuca plate. *Geophysical Research Letters* 30: 1095. http://dx.doi.org/10.1029/2002GL015511.

Chen W-P (1995) En echelon ruptures during the great Bolivian earthquake of 1994. *Geophysical Research Letters* 22: 2261–2264.

Chen P-F, Bina CR, and Okal EA (2004) A global survey of stress orientations in subducting slabs as revealed by intermediate-depth earthquakes. *Geophysical Journal International* 159: 721–733. http://dx.doi.org/10.1111/j.1365-246x.2004.02450.x.

Chen W-P and Brudzinski MR (2001) Evidence for a large-scale remnant of subducted lithosphere beneath Fiji. *Science* 292: 2475–2479.

Choy GL, McGarr A, Kirby SH, and Boatwright J (2006) In: Abercrombie RE, McGarr A, Kanamori H, and Di Toro G (eds.) *Earthquakes: Radiated Energy and the Physics of Faulting. An Overview of the Global Variability in Radiated Energy and Apparent Stress*, vol. 170, pp. 43–58. Washington, DC: American Geophysical Union.

Christova C and Scholz CH (2003) Stresses in the Vanuatu subducting slab: A test of two hypotheses. *Geophysical Research Letters* 30(15): 1790. http://dx.doi.org/10.1029/2003GL017701.

Comte D and Suárez G (1994) An inverted double seismic zone in Chile: Evidence of phase transformation in the subducted slab. *Science* 263: 212–215.

Davies JH (1999) The role of hydraulic fractures and intermediate depth earthquakes in subduction zone magmatism. *Nature* 398: 142–145.

Devaux JP, Fleitout L, Schubert G, and Anderson C (2000) Stresses in a subducting slab in the presence of a metastable olivine wedge. *Journal of Geophysical Research* 105: 13365–13373.

Devaux JP, Schubert G, and Anderson C (1997) Formation of a metastable olivine wedge in a descending slab. *Journal of Geophysical Research* 102: 24627–24637.

Dziewonski AM and Gilbert F (1974) Temporal variation of the seismic moment tensor and the evidence of precursive compression for two deep earthquakes. *Nature* 257: 185–188.

Estabrook CH (1999) Body wave inversion of the 1970 and 1963 South American large deep-focus earthquakes. *Journal of Geophysical Research* 104: 28751–28767.

Estabrook CH and Bock G (1995) Rupture history of the great Bolivian earthquake: Slab interaction with the 660-km discontinuity? *Geophysical Research Letters* 22: 2277–2280.

Fletcher RC and Pollard DD (1981) Anti-crack model for pressure solution surfaces. *Geology* 9: 419–424.

Frohlich C (1987) Aftershocks and temporal cluster of deep earthquakes. *Journal of Geophysical Research* 92: 13944–13957.

Frohlich C (1989) The nature of deep-focus earthquakes. *Annual Review of Earth and Planetary Sciences* 17: 227–254.

Frohlich C (1994) Earthquakes with non-double-couple focal mechanisms. *Science* 264: 804–809.

Frohlich C (1995) Characteristics of well-determined non-double-couple earthquakes in the Harvard CMT catalog. *Physics of the Earth and Planetary Interiors* 91: 213–228.

Frohlich C (2006a) *Deep Earthquakes*. Cambridge, UK: Cambridge University Press.

Frohlich C (2006b) A simple analytical method to calculate the thermal parameter and temperature within subducted lithosphere. *Physics of the Earth and Planetary Interiors* 155: 281–285.

Frohlich C and Davis SD (1993) Teleseismic b-values – Or, much ado about 1.0. *Journal of Geophysical Research* 98: 631–644.

Fukao Y (1972) Source process of a large deep earthquake and its tectonic implications: The western Brazil earthquake of 1963. *Physics of the Earth and Planetary Interiors* 5: 61–76.

Fukao Y and Kikuchi M (1987) Source retrieval for mantle earthquakes by iterative deconvolution of long-period P waves. *Tectonophysics* 144: 249–269.

Furumoto M (1977) Spatio-temporal history of the deep Colombian earthquake of 1970. *Physics of the Earth and Planetary Interiors* 15: 1–12.

Giardini D (1988) Frequency distribution and quantification of deep earthquakes. *Journal of Geophysical Research* 93: 2095–2105.

Giardini D (1992) Space–time distribution of deep seismic deformation in Tonga. *Physics of the Earth and Planetary Interiors* 74: 75–88.

Giardini D and Lundgren P (1995) The June 9 Bolivia and March 9 Fiji deep earthquakes of 1994: II. Geodynamic implications. *Geophysical Research Letters* 22: 2281–2284.

Giardini D and Woodhouse JH (1984) Deep seismicity and modes of deformation in Tonga subduction zone. *Nature* 307: 505–509. http://dx.doi.org/10.1038/307505a0.

Glennon MA and Chen W-P (1995) Ruptures of deep-focus earthquakes in the northwestern pacific and their implications on seismogenesis. *Geophysical Journal International* 120: 706–720.

Gorbatov A and Kostoglodov V (1997) Maximum depth of seismicity and thermal parameter of the subducting slab: General empirical relation and its application. *Tectonophysics* 277: 165–187.

Gorbatov A, Suarez G, Kostoglodov V, and Gordeev E (1994) A double-planed seismic zone in Kamchatka from local and teleseismic data. *Geophysical Research Letters* 21: 1675–1678.

Goto K, Suzuki Z, and Hamaguchi H (1987) Stress distribution due to olivine–spinel phase transformation in descending plate and deep focus earthquakes. *Journal of Geophysical Research* 92: 13811–13820.

Green HW and Burnley PC (1989) A new self-organizing mechanism for deep-focus earthquakes. *Nature* 341: 733–737.

Green HW and Houston H (1995) The mechanics of deep earthquakes. *Annual Review of Earth and Planetary Sciences* 23: 169–213.

Green HW and Marone C (2002) Instability of deformation. In: Karato S and Wenk H (eds.) *Reviews in Mineralogy and Geochemistry, Plasticity of Minerals and Rocks*, vol. 51, pp. 181–1999. Oxford, UK: Blackwell.

Green HW and Zhou Y (1996) Transformation-induced faulting requires an exothermic reaction and explains the cessation of earthquakes at the base of the mantle transition zone. *Tectonophysics* 256: 39–56.

Griggs D and Handin J (1960) Observations on fracture and a hypothesis of earthquakes. *Geological Society of America Memoir* 79: 347–373.

Guest A, Schubert G, and Gable CW (2003) Stress field in the subducting lithosphere and comparison with deep earthquakes in Tonga. *Journal of Geophysical Research* 108: B6. http://dx.doi.org/10.1029/2002JB002161.

Guest A, Schubert G, and Gable CW (2004) Stresses along the metastable wedge of olivine in a subducting slab: Possible explanation for the Tonga double seismic layer. *Physics of the Earth and Planetary Interiors* 141: 253–267.

Hacker BR, Peacock SM, Abers GA, and Holloway SD (2003) Subduction factory 2. Are intermediate-depth earthquakes in subducting slabs linked to metamorphic dehydration reactions? *Journal of Geophysical Research* 108: 24627–24637. http://dx.doi.org/10.1029/2001JB001129.

Hara T, Kuge K, and Kawakatsu H (1995) Determination of the isotropic component of the 1994 Bolivia deep earthquake. *Geophysical Research Letters* 22: 2265–2268.

Hara T, Kuge K, and Kawakatsu H (1996) Determination of the isotropic component of deep focus earthquakes by inversion of normal-mode data. *Geophysical Journal International* 127: 515–528.

Hasegawa A, Horiuchi S, and Umino N (1994) Seismic structure of the northeastern Japan convergent margin: A synthesis. *Physics of the Earth and Planetary Interiors* 99(B11): 22295–22312.

Helffrich G and Brodholt J (1991) Relationship of deep seismicity to the thermal structure of subducted lithosphere. *Nature* 353: 252–255.

Hill DP, Reasenberg PA, Michael A, et al. (1993) Seismicity remotely triggered by the magnitude 7.3 Landers California Earthquake. *Science* 260: 1617–1623.

Hobbs BE and Ord A (1988) Plastic instabilities: Implications for the origin of intermediate and deep focus earthquakes. *Journal of Geophysical Research* 93: 10521–10540.

Hogrefe A, Rubie DC, Sharp TG, and Seifert F (1994) Metastability of enstatite in deep subducting lithosphere. *Nature* 372: 351–353.

Houston H (2001) Influence of depth, focal mechanism, and tectonic setting on the shape and duration of earthquake source time functions. *Journal of Geophysical Research* 106: 11137–11150.

Houston H, Benz HM, and Vidale JE (1998) Time functions of deep earthquakes from broadband and short-period stacks. *Journal of Geophysical Research* 103: 29895–29913.

Houston H and Vidale JE (1994) Temporal distribution of seismic radiation during deep earthquake rupture. *Science* 265: 771–774.

Houston H and Williams Q (1991) Fast rise times and the physical mechanism of deep earthquakes. *Nature* 352: 520–522.

Huang W-C and Okal EA (1998) Centroid moment tensor solutions for deep earthquakes predating the digital era: Discussion and inferences. *Physics of the Earth and Planetary Interiors* 106: 191–218. http://dx.doi.org/10.1016/S0031-9201(97)00111-8.

Igarashi T, Matsuzawa T, Umino N, and Hasegawa A (2001) Spatial distribution of focal mechanisms for interplate and intraplate earthquakes associated with the subducting Pacific plate beneath the northeastern Japan arc: A tripled-planed deep seismic zone. *Journal of Geophysical Research* 106: 2177–2191.

Ihmle PF (1998) On the interpretation of subevents in teleseismic waveforms: The 1994 Bolivia deep earthquake revisited. *Journal of Geophysical Research* 103: 17919–17932.

Iidaka T and Furukawa Y (1994) Double seismic zone for deep earthquakes in the Izu–Bonin subduction zone. *Science* 263(5150): 1116–1118.

Iidaka T and Suetsugu D (1992) Seismological evidence for metastable olivine inside a subducting slab. *Nature* 356: 593–595. http://dx.doi.org/10.1038/356593a0.

Isacks BL and Barazangi M (1977) Geometry of Benioff zones: Lateral segmentation and downwards bending of the subducted lithosphere. In: Talwani M and Pitman WC III (eds.) *Island Arcs, Deep Sea Trenches and Back-Arc Basins*, pp. 99–114. Washington, DC: AGU.

Isacks B and Molnar P (1971) Distribution of stresses in the descending lithosphere from a global survey of focal mechanism solutions of mantle earthquakes. *Reviews of Geophysics* 9: 103–174.

Ito E and Sato H (1991) Aseismicity in the lower mantle by superplasticity of the descending slab. *Nature* 351: 140–141.

Jiang G, Zhao D, and Zhang Z (2008) Seismic evidence for a metastable olivine wedge in the subducting Pacific slab under Japan Sea. *Earth and Planetary Science Letters* 270: 300–307. http://dx.doi.org/10.1016/j.epsl.2008.03.037.

Kagan YY (1999) Universality of the seismic moment–frequency relation. *Pure and Applied Geophysics* 155: 537–573.

Kanamori H and Anderson DL (1975) Theoretical basis of some empirical relations in seismology. *Bulletin of the Seismological Society of America* 65: 1073–1095.

Kanamori H, Anderson DL, and Heaton TH (1998) Frictional melting during the rupture of the 1994 Bolivian earthquake. *Science* 279: 839–842.

Kaneshima S, Okamoto T, and Takenaka H (2007) Evidence for a metastable olivine wedge inside the subducted Mariana slab. *Earth and Planetary Science Letters* 258: 219–227. http://dx.doi.org/10.1016/j.epsl.2007.03.035.

Kao H and Chen W-P (1994) The double seismic zone in Kurile–Kamchatka – The tale of two overlapping zones. *Journal of Geophysical Research* 99: 6913–6930.

Karato S, Riedel MR, and Yuen DA (2001) Rheological structure and deformation of subducted slabs in the mantle transition zone: Implications for mantle circulation and deep earthquakes. *Physics of the Earth and Planetary Interiors* 127: 83–108.

Kawakatsu H (1986) Downdip tensional earthquakes beneath the Tonga arc: A double seismic zone. *Journal of Geophysical Research* 91: 6432–6440.

Kawakatsu H (1996) Observability of the isotropic component of a moment tensor. *Geophysical Journal International* 126: 525–544.

Kawakatsu H and Yoshioka S (2011) Metastable Olivine wedge and deep dry cold slab beneath southwest Japan. *Earth and Planetary Science Letters* 303: 1–10. http://dx.doi.org/10.1016/j.epsl.2011.01.008.

Kelemen PB and Hirth G (2007) Aperiodic shear-heating mechanism for intermediate-depth earthquakes in the mantle. *Nature* 446: 787–790. http://dx.doi.org/10.1038/nature05717.

Kikuchi M and Kanamori H (1994) The mechanism of the deep Bolivia earthquake of June 9, 1994. *Geophysical Research Letters* 21: 2341–2344.

Kirby SH, Durham WB, and Stern L (1991) Mantle phase changes and deep earthquake faulting in subducting lithosphere. *Science* 252: 216–225.

Kirby SH, Stein S, Okal EA, and Rubie DC (1996) Metastable mantle phase transformations and deep earthquakes in subducting oceanic lithosphere. *Reviews of Geophysics and Space Physics* 34: 261–306.

Kiser E, Ishii M, Langmuir CH, Shearer PM, and Hirose H (2011) Insights into the mechanism of intermediate-depth earthquakes from source properties as imaged by back projection of multiple seismic phases. *Journal of Geophysical Research* 116: B06310. http://dx.doi.org/10.1029/2010JB007831.

Kita S, Okada T, Hasegawa A, Nakajima J, and Matsuzawa T (2010) Existence of interplane earthquakes and neutral stress boundary between the upper and lower planes of the double seismic zone beneath Tohoku and Hokkaido, northeastern Japan. *Tectonophysics* 496: 68–82. http://dx.doi.org/10.1016/j.tecto.2010.10.010, L24310.

Kita S, Okada T, Nakajima J, Matsuzawa T, and Hasegawa A (2006) Existence of a seismic belt in the upper plane of the double seismic zone extending in the along-arc direction at depths of 70–100 km beneath NE Japan. *Geophysical Research Letters* 33: L24310. http://dx.doi.org/10.1029/2006GL028239.

Kohlstedt DL, Keppler H, and Rubie DC (1996) Solubility of water in the α, β and γ-phases of $(Mg, Fe)_2SiO_4$. *Contributions to Mineralogy and Petrology* 123: 345–357.

Koper KD and Wiens DA (2000) The waveguide effect of metastable olivine in slabs. *Geophysical Research Letters* 27: 581–584.

Kuge K (1994) Rapid rupture and complex faulting of the May 12, 1990, Sakhalin deep earthquake – Analysis of regional and teleseismic broad-band data. *Journal of Geophysical Research* 99: 2671–2685.

Lockner DA, Byerlee JD, Kuksenko V, Ponomarev A, and Sidorin A (1991) Quasi-static fault growth and shear fracture energy in granite. *Nature* 350: 39–42.

Lundgren PR and Giardini D (1992) Seismicity, shear failure and modes of deformation in deep subduction zones. *Physics of the Earth and Planetary Interiors* 74: 63–74.

Lundgren P and Giardini D (1994) Seismicity, shear failure and modes of deformation in deep subduction zones. *Journal of Geophysical Research* 99: 15833–15842.

Lundgren P and Giardini D (1995) The June 9 Bolivia and March 9 Fiji deep earthquakes of 1994. Part I: Source processes. *Geophysical Research Letters* 22: 2241–2244.

McGuire JJ, Wiens D, Shore P, and Bevis M (1997) The March 9, 1994 (Mw 7.6) deep Tonga earthquake: Rupture outside the seismically active slab. *Geophysical Research Letters* 102: 15163–15182.

Morris J and Tera F (1989) 10 Be and 9 Be in mineral separates and whole rocks from volcanic arcs: Implications for sediment subduction. *Geochemica et Cosmochimica Acta* 53: 3197–3206.

Mosenfelder JL, Marton FC, Ross CR, Kerschhofer L, and Rubie DC (2001) Experimental constraints on the depth of olivine metastability in subducting lithosphere. *Physics of the Earth and Planetary Interiors* 127(1–4): 165–180.

Myers SC, Wallace TC, Beck SL, et al. (1995) Implications of spatial and temporal development of the aftershock sequence for the M_w 8.3 June 9, 1994 deep Bolivian earthquake. *Geophysical Research Letters* 22: 2269–2272.

Myhill R and Warren LM (2012) Fault plane orientations of deep earthquakes in the Izu–Bonin–Marianas subduction zone. *Journal of Geophysical Research* 117: B06307. http://dx.doi.org/10.1029/2011JB009047.

Nyffenegger P and Frohlich C (2000) Aftershock occurrence rate decay properties for intermediate and deep earthquake sequences. *Geophysical Research Letters* 27: 1215–1218.

Ogawa M (1987) Shear instability in a viscoelastic material as the cause for deep focus earthquakes. *Journal of Geophysical Research* 92: 13801–13810.

Okal EA (2001) Detached' deep earthquakes: Are they really? *Physics of the Earth and Planetary Interiors* 127: 109–143.

Okal EA and Kirby SH (1998) Deep earthquakes beneath the Fiji basin, SW Pacific: Earth's most intense deep seismicity in stagnant slabs. *Physics of the Earth and Planetary Interiors* 109: 25–63.

Omori S, Kamiya S, Maruyama S, and Zhao D (2002) Morphology of the intraslab seismic zone and devolatilization phase equilibria of the subducting slab peridotite. *Bulletin of the Earthquake Research Institute* 76: 455–478.

Omori S, Komabayashi T, and Maruyama S (2004) Dehydration and earthquakes in the subducting slab: Empirical link in intermediate and deep seismic zones. *Physics of the Earth and Planetary Interiors* 146: 297–311. http://dx.doi.org/10.1016/j.pepi.2003.08.014.

Pardo M and Suarez G (1995) Shape of the subducted Rivera and Cocos plates in southern Mexico: Seismic and tectonic implications. *Journal of Geophysical Research* 100: 12357–12373.

Peacock SM (2001) Are the lower planes of double seismic zones caused by serpentine dehydration in subducting oceanic mantle? *Geology* 29: 299–302.

Persh SE and Houston H (2004a) Deep earthquake rupture histories determined by global stacking of broadband P waveforms. *Journal of Geophysical Research* 109: 4311. http://dx.doi.org/10.1029/2003JB002762.

Persh SE and Houston H (2004b) Strongly depth-dependent aftershock production in deep earthquakes. *Bulletin of the Seismological Society of America* 94: 1808–1816.

Petit JP and Barquins M (1988) Can natural faults propagate under Mode II conditions. *Tectonics* 7: 1243–1256.

Ranero CR, Phipps-Morgan J, McIntosh K, and Reichert C (2003) Bending-related faulting and mantle serpentinization at the Middle America trench. *Nature* 425: 367–373.

Ranero CR, Villasenor A, Morgan JP, and Weinrebe W (2005) Relationship between bend–faulting at trenches and intermediate-depth seismicity. *Geochemistry, Geophysics, Geosystems* 6: Q12002.

Rayleigh CB and Paterson MS (1965) Experimental deformation of serpentinite and its tectonic implications. *Journal of Geophysical Research* 70: 3965–3985.

Rietbrock A and Waldhauser F (2004) A narrowly spaced double-seismic zone in the subducting Nazca plate. *Geophysical Research Letters* 31: L10608. http://dx.doi.org/10.1029/2004GL019610.

Robinson R (1986) Seismicity, structure and tectonics of the Wellington region, New Zealand. *Geophysical Journal of the Royal Astronomical Society* 87: 379–409.

Russakoff D, Ekstrom G, and Tromp J (1997) A new analysis of the 1970 Columbia earthquake and its isotropic component. *Journal of Geophysical Research* 102: 20423–20434.

Scholz CH (2002) *The Mechanics of Earthquakes and Faulting*. Cambridge, UK: Cambridge University Press.

Schubnel A, Brunet F, Hilairet N, Gasc J, Wang Y, and Green HW (2013) Deep-focus earthquake analogs recorded at high pressure and temperature in the laboratory. *Science* 341: 1377–1380. http://dx.doi.org/10.1126/science.1240206.

Silver PG, Beck SL, Wallace TC, et al. (1995) Rupture characteristics of the deep Bolivian earthquake of June 9, 1994 and the mechanism of deep-focus earthquakes. *Science* 268: 69–73.

Stein S and Wysession M (2003) *An Introduction to Seismology, Earthquakes, and Earth Structure*. Oxford: Blackwell Publishing Limited.

Suzuki M and Yagi Y (2011) Depth dependence of rupture velocity in deep earthquakes. *Geophysical Research Letters* 38: L05308. http://dx.doi.org/10.1029/2011GL046807.

Syracuse EM, van Keken PE, and Abers GA (2010) The global range of subduction zone thermal models. *Physics of the Earth and Planetary Interiors* 183: 73–90. http://dx.doi.org/10.1016/j.pepi.2010.02.004.

Thompson AB (1992) Water in the Earth's upper mantle. *Nature* 358: 295–302.

Tibi R, Bock G, and Estabrook CH (2002) Seismic body wave constraint on mechanisms of intermediate-depth earthquakes. *Journal of Geophysical Research* 107: 2047. http://dx.doi.org/10.1029/2001JB000361.

Tibi R, Bock B, and Wiens DA (2003a) Source characteristics of large deep earthquakes: Constraint on the faulting mechanism at great depths. *Journal of Geophysical Research* 108: 2091. http://dx.doi.org/10.1029/2002JB001948.

Tibi R, Estabrook CH, and Bock G (1999) The 1996 June 17 Flores Sea and 1994 March 9 Fiji–Tonga earthquakes: Source processes and deep earthquake mechanisms. *Geophysical Journal International* 138(3): 625–642. http://dx.doi.org/10.1046/j.1365-246x.1999.00879.x.

Tibi R, Wiens DA, and Hildebrand JA (2001) Aftershock locations and rupture characteristics of the 1995 Mariana deep earthquake. *Geophysical Research Letters* 28: 4311–4314.

Tibi R, Wiens DA, and Inoue H (2003b) Remote triggering of deep earthquakes in the 2002 Tonga sequence. *Nature* 424: 921–925.

Tocheport A, Rivera L, and Chevrot S (2006) A systematic study of source time functions and moment tensors of intermediate and deep earthquakes. *Journal of Geophysical Research* 112: B07311.

van Keken PE, Hacker BR, Syracuse EM, and Abers GA (2011) Depth-dependent flux of H_2O from subducting slabs worldwide. *Journal of Geophysical Research* 116: B01401. http://dx.doi.org/10.1029/2010JB007922.

Venkataraman A and Kanamori H (2004) Observational constraints on the fracture energy of subduction zone earthquakes. *Journal of Geophysical Research* 109: B05302. http://dx.doi.org/10.1029/2003JB002549.

Vidale JE and Benz HM (1992) Upper-mantle seismic discontinuities and the thermal structure of subduction zones. *Nature* 356: 678–683.

Vidale JE and Houston H (1993) The depth dependence of earthquake duration and implications for rupture mechanisms. *Nature* 365: 45–47.

Vidale JE, Williams Q, and Houston H (1991) Waveform effects of a metastable olivine tongue in subducting slabs. *Geophysical Research Letters* 18: 2201–2204.

Wang K (2002) Unbending combined with dehydration embrittlement as a cause for double and triple seismic zones. *Geophysical Research Letters* 29: 1889–1892. http://dx.doi.org/10.1029/2002GL015441.

Warren LM (2010) Measurement of differential rupture duration as constraints on the source finiteness of deep-focus earthquakes: 2. Synthetic tests to estimate errors in rupture vectors and their effect on fault plane identification. *Journal of Geophysical Research* 115: B09307.

Warren LM, Hughes AN, and Silver PG (2007) Earthquake mechanics and deformation in the Tonga–Kermadec subduction zone from fault-plane orientations of intermediate- and deep-focus earthquakes. *Journal of Geophysical Research* 112: B05314. http://dx.doi.org/10.1029/2006JB004677.

Warren LM, Langstaff MA, and Siver PG (2008) Fault plane orientations of intermediate-depth earthquakes in the Middle America Trench. *Journal of Geophysical Research* 113: B10304. http://dx.doi.org/10.1029/2007JB005028.

Wiens DA (2001) Seismological constraints on the mechanisms of deep earthquakes: Temperature dependence of deep earthquake source properties. *Physics of the Earth and Planetary Interiors* 127: 145–163.

Wiens DA and Gilbert HJ (1996) Effect of slab temperature on deep-earthquakes aftershock productivity and magnitude–frequency relations. *Nature* 384: 153–156.

Wiens DA and McGuire JJ (1995) The 1994 Bolivia and Tonga events – Fundamentally different types of deep earthquakes. *Geophysical Research Letters* 22: 2245–2248.

Wiens DA and McGuire J (2000) Aftershocks of the March 9, 1994 Tonga earthquake: The strongest known deep aftershock sequence. *Journal of Geophysical Research* 105: 19067–19083.

Wiens DA, McGuire JJ, and Shore PJ (1993) Evidence for transformational faulting from a deep double seismic zone in Tonga. *Nature* 364(6440): 790–793.

Wiens DA and Snider NO (2001) Repeating deep earthquakes: Evidence for fault reactivation at great depth. *Science* 293(5534): 1463–1466.

Winslow NW and Ruff LJ (1999) A hybrid method for calculating the radiated wave energy of deep earthquakes. *Physics of the Earth and Planetary Interiors* 115: 181–190.

Wu L-R and Chen W-P (2001) Rupture of the large (Mw 7.8), deep earthquake of 1973 beneath the Japan Sea with implications for seismogenesis. *Bulletin of the Seismological Society of America* 91: 102–111.

Yamasaki T and Seno T (2003) Double seismic zone and dehydration embrittlement of the subducting slab. *Journal of Geophysical Research* 108: 2212. http://dx.doi.org/10.1029/2002JB001918.

Ye L, Lay T, Kanamori H, and Koper KD (2013) Energy Release of the 2013 M_w 8.3 Sea of Okhotsk Earthquake and Deep Slab Stress Heterogeneity. *Science* 341: 1380–1384. http://dx.doi.org/10.1126/science.1242032.

Yoshioka S, Daessler R, and Yuen DA (1997) Stress fields associated with metastable phase transitions in descending slabs and deep-focus earthquakes. *Physics of the Earth and Planetary Interiors* 104: 345–361.

Zhan Z, Kanamori H, Tsai VC, Helmberger DV, and Wei S (2014) Rupture complexity of the 1994 Bolivia and 2013 Sea of Okhotsk deep earthquakes. *Earth and Planetary Science Letters* 385: 89–96.

Relevant Websites

http://www.sionizconter.ucsd.edu – SIO Visualization Center at Scripps Institute of Oceanography, Institute for Geophysics and Planetary Physics (this website will give you full directions on how to download and use the iView3D software).

4.14 Volcanology 101

CG Newhall, Mirisbiris Garden and Nature Center, Sto. Domingo, Albay, Philippines

© 2015 Elsevier B.V. All rights reserved.

4.14.1	**Introduction to Volcanic Systems**	357
4.14.2	**Magma and Gas**	357
4.14.2.1	Melt	357
4.14.2.2	Crystals	357
4.14.2.3	Volatile Phases: Both Dissolved and Exsolved	358
4.14.2.3.1	Gas supply	358
4.14.2.3.2	Degassing	358
4.14.2.3.3	Gas budgets	359
4.14.2.4	Viscosity, Density, and Volatile Content	361
4.14.3	**Intrusions and Convection**	363
4.14.3.1	Orientation, Geometry, and Dimensions of Intrusions	363
4.14.3.2	Drivers of One-Way Magma Ascent	363
4.14.3.3	Drivers of Convective Magma Ascent with Return Flow	363
4.14.3.4	Magma Mixing: Conditions, Evidence, and Effects	364
4.14.3.5	Geospeedometers of Magma Ascent, Convection, and Mixing	365
4.14.4	**Pressurization Within and Around Magma**	366
4.14.4.1	Increase in Overpressure	366
4.14.4.1.1	Volatile accumulation	366
4.14.4.1.2	Magma ascent and volatile expansion	366
4.14.4.1.3	Heating and/or compression of groundwater	366
4.14.4.1.4	Lithostatic pressure	366
4.14.4.1.5	Regional tectonics: quasistatic and dynamic strain	367
4.14.4.2	Decrease in Overpressure	367
4.14.4.2.1	Gas leaks	367
4.14.4.2.2	Eruptions	367
4.14.4.2.3	Groundwater diffusion and dissipation of pore water pressures; Breakage of self-seals	367
4.14.4.3	Dynamic Balance between Pressurization and Depressurization	367
4.14.5	**Volcanic Unrest**	367
4.14.5.1	Gas Emissions	367
4.14.5.2	Seismicity	368
4.14.5.3	Ground Deformation	369
4.14.5.4	Other Changes: Gravitational, Magnetic, Electrical, Thermal, and Hydrologic	370
4.14.6	**Volcanic Eruptions**	371
4.14.6.1	Eruption Triggering Mechanisms	371
4.14.6.1.1	Internal	371
4.14.6.1.2	External	371
4.14.6.2	Explosive versus Effusive Eruptions: Fragmentation of Magma	372
4.14.6.3	Measures of Magnitude and Intensity (Volume Dense Rock Equivalent, Mass Discharge Rate, and VEI)	373
4.14.6.4	Eruptive Phenomena	374
4.14.6.4.1	Lava flows and lava domes	374
4.14.6.4.2	Tephra fall	375
4.14.6.4.3	Pyroclastic flows and surges	375
4.14.7	**Construction and Erosion of Volcanoes**	375
4.14.7.1	Hydrothermal Alteration	376
4.14.7.2	Sector Collapses	376
4.14.7.3	Concentrated Sediment–Water Flows (Lahars)	376
4.14.7.4	Normal Streamflow, Sediment Issues, and Flooding	377
4.14.8	**People and Volcanoes**	377
4.14.8.1	Risk Mitigation	377
4.14.8.2	Eruption Forecasts	377
4.14.8.2.1	Long-range forecasts	377

4.14.8.2.2 Short-range forecasts 377
4.14.8.2.3 Successful forecasts and false alarms 378
4.14.8.3 Quantification and Acceptance of Risk 378
4.14.9 Future Directions 379
4.14.10 Conclusion 380
Acknowledgment 381
References 381

Glossary

Degassing (a) Exsolution of volatiles from melt into volatilerich bubbles that are still contained within magma or (b) escape of such bubbles from magma into country rock, a hydrothermal system, or into the atmosphere. In this chapter, we distinguish these as melt degassing versus magma degassing.

Dike A tabular igneous intrusion that cuts across planar structures of surrounding rock, often cutting vertically through horizontal strata.

Eruption Extrusion or explosive ejection of volcanic material onto the Earth's surface or into the atmosphere – (1) phreatic eruption: steam blast, without new magma; (2) magmatic eruption: eruption predominantly of magma; and (3) phreatomagmatic eruption: hybrid of the two. Many eruptions progress from phreatic, to phreatomagmatic, to magmatic, and some revert back to phreatomagmatic in their late stages as groundwater regains access to magma.

Felsic Adjective describing minerals and rocks rich in silica and aluminum and poor in iron and magnesium. Felsic rocks include dacite, rhyolite, their intrusive equivalents of granodiorite and granite, and a variety of less common rock types with higher or lower contents of Na and K.

Groundmass Fine-grained mixture of glass and microlites, forming the interstitial matrix of volcanic rocks containing phenocrysts.

Mafic Adjective describing minerals and rocks rich in iron and magnesium and relatively poor in silica and aluminum. Mafic rocks include basalt, its intrusive equivalent of gabbro, and a variety of less common rock types with higher or lower contents of Na and K. Andesite and its intrusive equivalent, diorite, with slightly higher silica content, are intermediate between mafic and felsic.

Magma Molten rock, consisting of up to three phases – melt, crystals, and exsolved volatiles (bubbles).

Magma mixing (and mingling) Mixing is mechanical and chemical homogenization of two magmas. Mingling is incomplete mixing and is seen in 'marble-cake' banded rocks.

Microlite Very small crystal in volcanic rock, typically $<<1$ mm in longest dimension, formed in volcanic conduits and lavas, shortly before, during, and shortly after eruption.

Oxygen fugacity (abbreviated f_o) A measure of the oxidizing or reducing character of a fluid. More precisely, the chemical activity of oxygen in a chemical solution, that is, the amount of oxygen available for chemical reactions. Typically expressed in log units, absolute or relative to the amount of free oxygen associated with an assemblage of minerals, for example, QFM, quartz (SiO_2)–fayalite (Fe_2SiO_2)–magnetite (Fe_3O_4).

Petrologic method The estimation of preeruption volatile contents from the highest volatile concentrations trapped in glass inclusions within phenocrysts and of degassing based on the difference between volatile concentrations in glass inclusions minus concentrations in matrix (groundmass) glass. A variation of the petrologic method approximates the concentration of volatiles in magma as 100 wt.% minus the analytic total of nonvolatile components.

Phenocryst Crystal in igneous rock that is notably larger than the average grain size of that rock, probably because it formed over extended periods within magma reservoirs several kilometers or more beneath the surface. No specific size definition, but almost always >1 mm in length.

Pumice Rapidly quenched ('frozen') magma foam. Low density and full of vesicles (bubble holes). Density often <1 g cm^{-3} so it floats in water, in contrast to nonvesicular rock densities of approximately 2.5 g cm^{-3}.

Scoria ('cinder') Volcanic rock with high vesicularity but less than that of pumice. Although there is a tendency for scoria to be mafic and pumice felsic, composition is not formally a part of the definition of pumice or scoria.

Second boiling Process whereby crystallization concentrates volatiles in the remaining melt phase and eventually causes them to exsolve into bubbles, mimicking a boiling process even though temperature may be constant or decreasing.

Sill A tabular igneous intrusion that parallels planar structures of surrounding rock, such as horizontal strata.

Ultramafic rocks Rocks containing low-silica contents and crystallizing large percentages of mafic minerals, for example, olivine, pyroxenes, and amphiboles. These include peridotite, eclogite, and a variety of less common combinations of mafic minerals.

Unrest An anomalous change from the background level of seismicity, ground deformation, gas emission, thermal output, and other measured parameters of a volcano. Some unrest leads to eruptions but much does not, so we use the term 'unrest' rather than 'precursor' unless the temporal or causal association with an eruption is clear. Eruptions are included within the definition of unrest.

Volcano A vent or cluster of vents from which molten or solid debris is erupted. Most volcanoes build edifices from erupted lava and fragmental debris; some that erupt very large volumes of magma create depressions in the ground, as in the case of calderas.

4.14.1 Introduction to Volcanic Systems

Active volcanoes are dynamic systems that can be instructively compared to active fault zones. Both are subject to stress and strain, both produce distinctive patterns of seismicity, and, in a sense, both exhibit stick–slip behavior with similar power-law magnitude–frequency relations. The distribution of areas at risk from active volcanoes and fault zones can be mapped and future hazards forecast over decades, centuries, or longer. However, beyond the similarities, volcanic systems involve high-temperature silicate melts, significant volumes of dissolved and undissolved volatiles, and important mass transfer before eruptions. Volcanic eruptions have far more pronounced precursors than do tectonic earthquakes, and, accordingly, in addition to long-term forecasts, both hour–day and week–month scale forecasts are often possible. Some of the precursors to eruptions are similar to those sought for earthquakes (e.g., precursory creep and microseismicity), while others are unique on account of the magma (e.g., leaks of high-temperature gas).

Some volcanoes are sensitive to static or dynamic tectonic strain, for example, Long Valley and Yellowstone Calderas (Hill, 2006; Hill et al., 2002), and, conversely, magmatic intrusions and hydrothermal pressurization can induce nearby tectonic seismicity (Legrand et al., 2002; White and Power, 2001).

This chapter describes the fundamental matter of volcanoes – molten rock (magma), which is composed of melt (typically silicate melt), crystals (dominantly, feldspars, quartz, and Fe–Mg silicates), and volatiles (mostly water). The volatile species are dissolved or occur as bubbles in the melt. Although magma is generated in the upper mantle and lower crust, this chapter is concerned mainly with its ascent from relatively shallow crustal reservoirs to the surface and associated pressurization of magma and surrounding groundwater. Volcanic unrest – volcanic seismicity, ground deformation, gas emission, and other changes – reflects intrusion, pressurization, and preeruption degassing of magma and forms the basis for short-term eruption forecasts. Finally, the chapter notes briefly different eruptive phenomena, posteruption erosion of volcanoes, and future directions in volcanology.

Volcanic eruptions range from a few cubic meters to thousands of cubic kilometers of magma, minutes to decades in duration, and gently effusive to astoundingly explosive. Perhaps, the most important control on the explosivity of an eruption appears to be the 'leakiness' of the system, that is, how much of its magma gas has leaked out since the previous eruption and before soon-to-erupt magma reaches the surface.

Worldwide, more than 500 million people live within zones that could be seriously affected by volcanic eruptions (Ewert and Harpel, 2004; Tilling, 1991). Millions more fly over or downwind from volcanoes every year. The average annual death toll from volcanic activity between 1800 and 2000 was $\sim$1000 year^{-1}. That average is dominated by just a few particularly sad disasters, such as the eruptions of Montagne Pelée in 1902 and of Nevado del Ruiz in 1985. During the past four decades, roughly 2 million people on the ground have been evacuated out of harm's way, and several million air passengers have been diverted away from dangerous ash clouds. Populations living on and near volcanoes are ever increasing, and volcanologists today are faced with rising expectations for accurate and timely eruption forecasts to keep these populations safe.

This chapter describes the author's views on the way volcanoes work. Volcanologists are strongly influenced by the volcanoes they have known. This chapter should be considered as a set of working hypotheses, judged likely today but subject to revision tomorrow. What follows is by no means comprehensive. Sigurdsson et al. (2000), Francis and Oppenheimer (2004), Schmincke (2004), Lockwood and Hazlett (2010), and Cashman and Sparks (2013) are excellent references on many more aspects of volcanology, including volcanic rocks and features, and Dobran (2001) gives mathematical descriptions of many volcanic processes.

4.14.2 Magma and Gas

4.14.2.1 Melt

Molten silicate, or in rare cases molten carbonate, is generated in the upper mantle beneath spreading centers and hot spots, in mantle wedges of arcs, and within the crust. Most mantle melts are basaltic in composition, commonly distinguished by their variable alkali (Na and K) contents. They originate by variable degrees (usually $<$20% partial melting of mantle peridotite). Crustal melts tend to be much more silicic, rhyolite or dacite, produced by $<$20% and often $<$10% partial melting of (mostly lower) crust or, in some cases, by fractionation of mantle-derived basalt. The percentage of partial melt present in a source region at any given time is considerably less than 20%, probably not more than a few percent. Mafic melts can have temperatures up to 1200 °C; rare ultramafic melts can be up to 1400 °C. Silicic melts can exist down to about 700 °C.

Experimental research on rapid solidification of melt into glass is described by Dingwell (1998, 2003).

The term magma (molten rock) encompasses not only the molten silicate but also any crystals that have formed or remain from incomplete melting, and volatiles that are either dissolved within the silicate melt or occur as bubbles in that melt (**Figure 1**).

Magmas can further differentiate by fractional crystallization in which early crystallizing phases and remaining melt are separated by crystal-settling and/or by filter-pressing melt out of crystal-rich mush, by assimilation (digestion) of country rock, and by magma mixing.

4.14.2.2 Crystals

Crystals form in melt as it cools or as it loses dissolved volatiles. Decreasing temperature and loss of dissolved water lead to polymerization of silica tetrahedra (SiO_4) in the melt that are the building blocks of common volcanic minerals like quartz (SiO_2), plagioclase feldspars (Ca/Na–Al silicates), alkali feldspars (Na/K–Al silicate), olivine (Mg/Fe silicate), pyroxenes (Ca/Mg/Fe silicates), hornblende and other amphiboles (hydrous Ca/Na/K–Mg/Fe/Al silicates), and biotite and other micas (hydrous K/Na–Mg/Fe/Al/Ti silicates) (slash indicates solid solution between these end-members, e.g., between Ca and Na in plagioclase feldspar). Other important volcanic minerals, small in volume but very useful in tracking petrologic

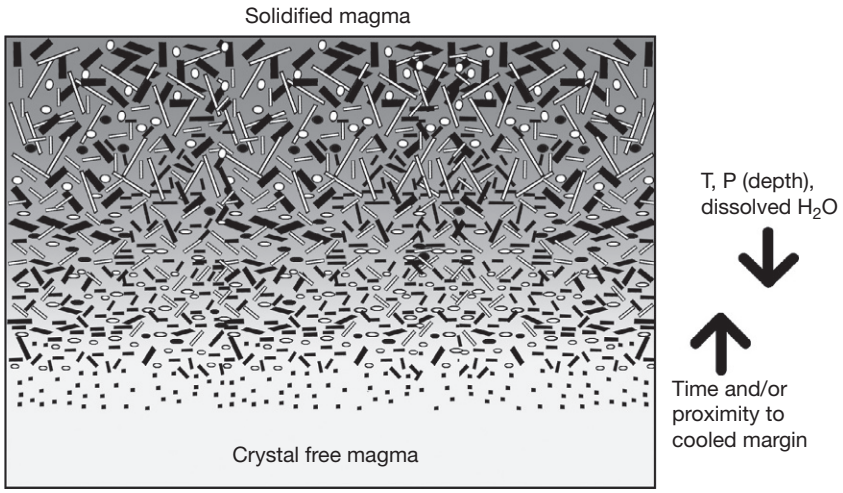


Figure 1 Components and evolution of magma, from melt to solid. Initially, at high temperature and pressure, only melt is present. As temperature declines, crystals will begin to form. As lithostatic pressure declines, water and other volatiles will also begin to exsolve into bubbles, inducing further crystallization. The sequence shown here from bottom to top, symbolizing change as magma moves from depth to the surface, is also repeated from center to margin of dikes and other intrusions. Reproduced from Marsh BD (2000) Magma chambers. In: Sigurdsson H, Houghton BF, McNutt SR, Rymer H, and Stix J (eds.) *Encyclopedia of Volcanoes*, pp. 191–206. San Diego, CA: Academic Press.

processes, are the Fe–Ti oxide minerals, which can be used to determine the temperature and oxygen fugacity of magma.

Crystals occur in two broad size groups, phenocrysts (typically, 1–10 mm) and groundmass crystals (microlites, typically $<<1$ mm). Phenocrysts form while magma is still in its reservoir, whereas microlites grow during final ascent and eruption of magma.

Volcanic rocks rarely contain more than 60% phenocrysts (Marsh, 1981). The reason is that higher phenocryst contents make magmas too viscous to erupt. The balance, 40%, is melt in the reservoir but can grow a significant percentage of microlites shortly before and during eruptions.

4.14.2.3 Volatile Phases: Both Dissolved and Exsolved

4.14.2.3.1 Gas supply

Volatile components of magma, dominantly H_2O, CO_2, sulfur gases SO_2 and H_2S, halogen gases HCl HF and HBr, hydrogen, and helium in decreasing order of abundance, are incorporated into magma during the initial melting of mantle or crust. Water content in magma can range from <0.1 to 8 wt.%. Carbon dioxide content can be as high as 1 wt.% (10 000 ppm) but is usually a fraction of this. Sulfur concentrations (expressed as elemental S) are typically 10–1000 ppm and occasionally several times higher. Beneath arcs, additional water and other volatiles come from subducted sediment and from metamorphism of downgoing, hydrothermally altered oceanic crust to eclogite. That water is released upward into the mantle wedge, where it fluxes melting and gets incorporated into magmas.

4.14.2.3.2 Degassing

There are limits to volatile solubility in silicate melts. The limits depend mostly strongly on pressure; the composition, temperature, and oxygen fugacity of the silicate melt; and the specific volatile species in question. Volatiles will form bubbles if they are added to or concentrated in melt in excess of their solubility. One way to add volatiles to magma is by convective and diffusive transfer from fresh magma entering the base of a magma reservoir. Another mechanism for volatile concentration is called 'second boiling,' in which volatiles, excluded from most crystal structures, are concentrated by magma crystallization into the remaining melt fraction and eventually exceed their solubilities in that melt. Temperature may be constant or even falling, yet volatiles are 'boiling,' exsolving from the melt into bubbles.

Bubbles will be of lower density than their host melt and can rise buoyantly through that melt and magma. Large bubbles (sometimes, up to several meters in diameter) are seen to rise up through fluid magma in Hawaii, at Mount Erebus in Antarctica, and elsewhere. This process ('separated flow') is an effective way to degas fluid magmas (Vergniolle and Jaupart, 1986). An everyday analogy is the bottle of soda pop or champagne from which bubbles rise freely once the confining (lithostatic) pressure is released. Eventually, the soda pop and magma 'go flat' unless there is resupply of volatiles from below. Degassing in viscous magmas is not as well understood. Clearly gases do escape – we can measure them at the surface and we also know that silicic magmas can 'go flat.' However, in viscous magmas, simple Stokes' law bubble rise is too slow to explain observed gas fluxes. Three alternate processes may be at play. One is convective transport of gas-rich magma up to a level at which it can foam, release its gas, and then sink in return flow as dense, now-degassed silicate melt and crystals. One can observe this process directly in lava lakes, and it is inferred at many other volcanoes for which the gas flux far exceeds what could have been dissolved in the volume of magma erupted over that same period (e.g., Allard, 1977; Allard et al., 1994; Andres et al., 1991; Witter et al., 2004, 2005). A second mechanism by which gas can escape from viscous magma is by diffusion or bubble transport over relatively short distances to the edge of magma conduits, followed by escape into high-permeability breccias that form along the

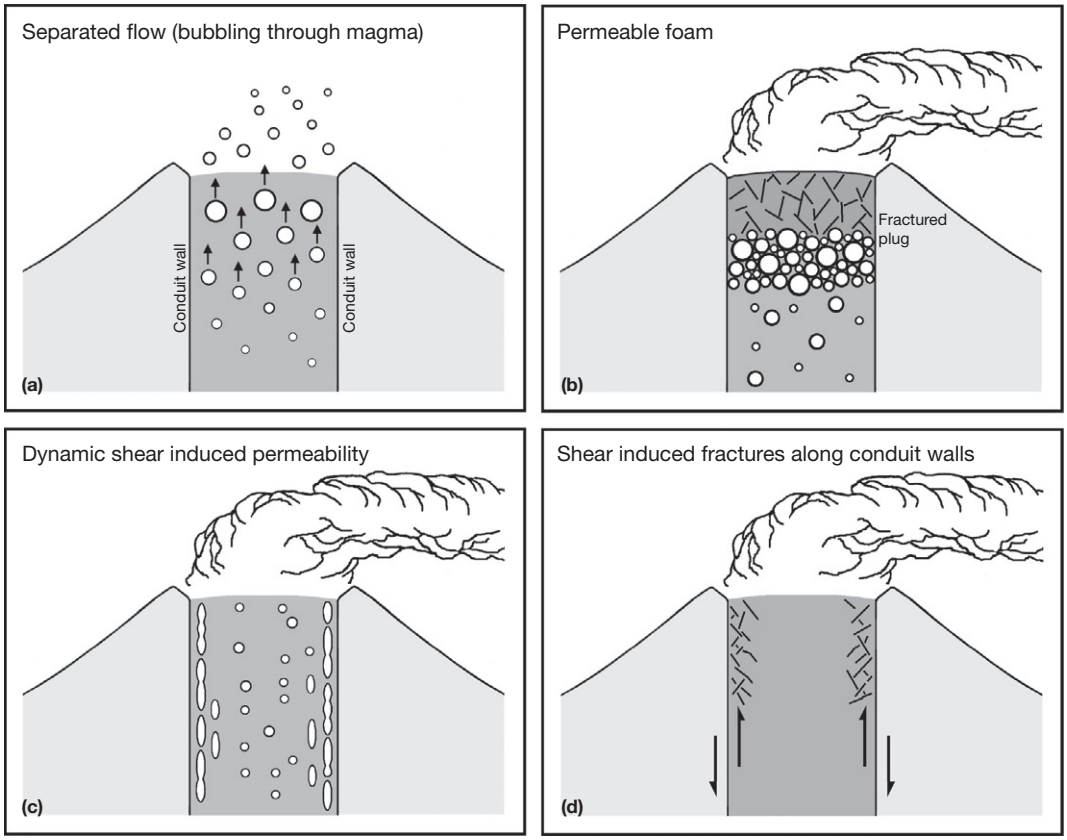


Figure 2 Various modes of gas permeability in magma: (a) separated flow of bubbles through melt, (b) development of a permeable foam at top of magma column, (c) dynamic shear-induced permeability (from bubbles and melt being stretched), and (d) shear-induced fractures along the margins of a volcanic conduit.

margins of some conduits by shear (Burgisser and Gardner, 2005; Jaupart and Allègre, 1991; Okumura et al., 2010; Rust et al., 2003; Stasiuk et al., 1996). Third, some dynamic bubble-to-bubble permeability may also develop within viscous magma as bubbles are brought into contact and interconnected by shear during magma ascent (Okumura et al., 2009). Various modes of degassing are illustrated in **Figure 2**.

An interesting paradox was further addressed by Okumura et al. (2013): How, if shear creates permeability, can volcanoes ever retain enough gas to drive highly explosive eruptions? The answer appears to be that shear is localized in a narrow ring along the margins of the conduit, allowing the rest of the magma to retain its gas and erupt explosively. Localization of extreme shear along the margin of a conduit is supported by a unique texture in pumice of the most explosive eruptions, in which phenocrysts were so intensely comminuted that the pumice appears to be lacking phenocrysts until one looks at the micron scale (Pallister et al., 1996; Polacci et al., 2001).

4.14.2.3.3 Gas budgets

The balance between (1) gas influx into crustal magma reservoirs (e.g., arriving dissolved in fresh basaltic magma), (2) continuous, passive degassing into the atmosphere and into hydrothermal systems, and (3) degassing during explosive eruptions is an important measure of the behavior of a volcano (**Figure 3**).

The gas influx term can be roughly estimated from a volume rate of magma supply (from the mantle and lower crust up into a shallow magma reservoir) times the concentration of each volatile species in that magma. A minimum estimate for volume rate of magma supply can come from the long-term eruption rate for that volcano; a better estimate that includes magma that does not erupt (intrusions) can come from long-term growth rate for a whole volcanic edifice or volcanic arc. If magma supply is relatively continuous, these long-term rates may also apply to shorter time intervals. Typical magma supply rates for arc volcanoes are 10^7 m^3 year^{-1}, plus or minus an order of magnitude; rates for ocean island and hot-spot volcanoes can be 10^8 m^3 year^{-1}, plus or minus an order of magnitude (Annen, 2009; Nakamura, 1974; Takada, 1999; White et al., 2006).

If one measures gases that are not common in the atmosphere, for example, SO_2, one can quantify passive degassing of magma between times of eruptions. Unfortunately, passive gas flux has been estimated consistently at only a few dozen volcanoes. Volcanoes instrumented by the NOVAC project (Galle et al., 2010) have consistent, continuous ground-based measures of SO_2. Elsewhere, measurements have been too infrequent to give reliable estimates. A promising development for systematic, space-based monitoring of passive degassing of SO_2, down to levels of hundreds of tonnes day^{-1}, is the new ozone mapping instrument (OMI), which began collecting

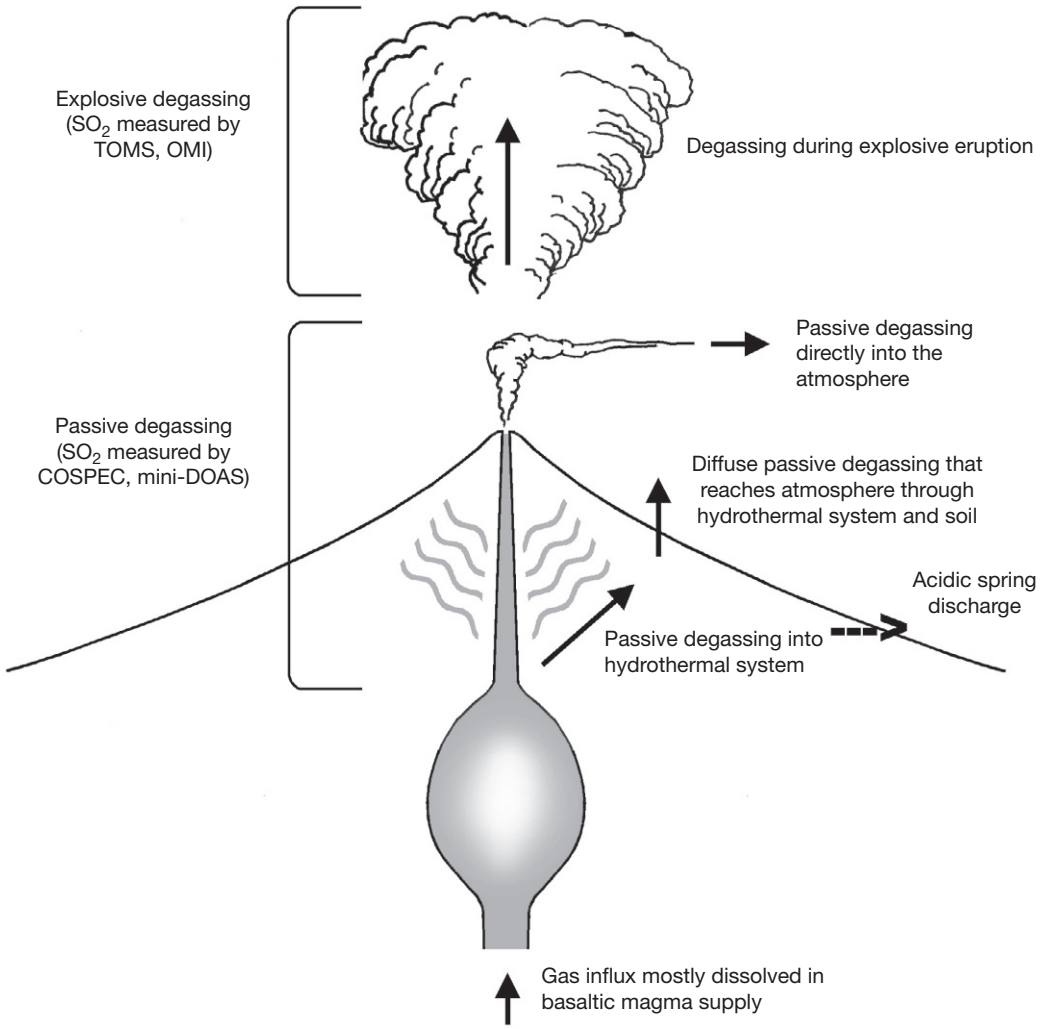


Figure 3 Elements of a simple volcanic gas budget, including input from depth, loss into a hydrothermal system, acidic spring discharge, passive loss to the atmosphere, and explosive loss to the atmosphere.

data in 2004 (Carn et al., 2009). Absorption ('scrubbing') of SO_2 into groundwater will prevent its release and measurement at the surface, so measurements of passive degassing are always minimum estimates.

A final term in a gas budget is the mass of gas released during eruptions. Again, using SO_2 (and estimating other species in relation to SO_2), the best data come from NASA's Total Ozone Mapping Spectrometer (TOMS) and now from OMI. Data from the largest eruptions have been gathered since 1977 (Bluth et al., 1997; Carn et al., 2003).

Of the three main terms in a gas budget – input, passive degassing, and eruptive degassing – the input term is the most poorly known. As a result, it is difficult to know the degree to which magma has already degassed before eruption. Yet, pre-eruption degassing may be the key to estimating explosive potential of an impending eruption.

One important insight that is appearing from gas budgets is that most magmas lose most of their gas passively, between eruptions. Degassing during explosive eruptions is spectacular, but less noticed, steady leakage dominates discharge. If we do a thought exercise by assuming concentrations of 1000 and 2500 ppm S in 10^7 m^3 $year^{-1}$ of fresh mafic magma supply to typically explosive arc volcanoes, multiply by the number of years since the last gas-depleting Volcanic Explosivity Index (VEI) ≥ 3 explosive eruption, convert from S to SO_2, and compare to the SO_2 release of the next VEI ≥ 3 eruption (Carn et al., 2003), we find that for most arc volcanoes, explosive gas release is <10% of inferred gas input. The difference must be released passively. This has interesting implications for conduit dynamics between eruptions as noted in **Section 4.14.3.3** and for explosive potential as noted in **Section 4.14.6.2**.

Volcanoes in Hawaii are classic examples of volcanoes that degas passively. Even within arcs, for the aforementioned thought exercise, there are sharp contrasts between volcanoes that release almost all of their gas passively (e.g., Mayon Volcano, **Figure 4(a)**) and those that retain a substantial percentage of their incoming gas and then release it in large explosive eruptions (e.g., Mount Pinatubo, **Figure 4(b)**) (Newhall, 2004).

Another important insight that has come from gas budgets is that some magmas appear to accumulate volatiles well in excess of saturation, with the excess volatiles accumulating in discrete fluid bubbles (Gerlach et al., 1996; Keppler, 1999; Luhr et al., 1984; Wallace, 2001; Wallace and Anderson,

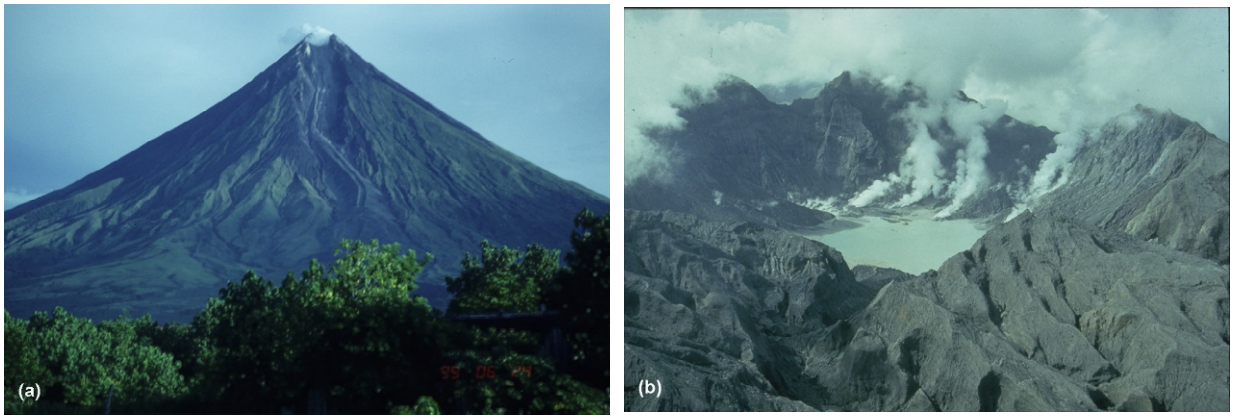


Figure 4 Two arc volcanoes with sharply different patterns of degassing. (a) Mayon Volcano, Philippines, which since 1978 has released ~95% of its gas passively and only ~5% explosively. Only 50 kt of SO_2 was recorded by the TOMS sensor, during the explosion of 1984. More sensitive ground-based monitoring showed about 150 kt of SO_2 released during the dominantly effusive eruption of 2006 and about 70 kt during the effusive 2009 eruption. Measured release during eruptions is, together, an order of magnitude less than what was measured in quiet periods before those eruptions. (b) Mount Pinatubo, Philippines, whose conduit was plugged for approximately 500 years so ~17 Mt of SO_2 (~60–70% of incoming SO_2) accumulated in the magma and was released in a large explosive eruption in 1991; the remaining ~30–40% was either released passively into the pre-1991 hydrothermal system or remains in the magma. (Assumption for both volcanoes: continuous deep supply of 10^7 m^3 basaltic magma per year, 1000 ppm S. TOMS data are from Bluth et al. (1997) and Carn et al. (2003).)

2000). Petrologic comparison of volatile contents in undegassed glass inclusions that were trapped long before eruptions in early formed phenocrysts, versus degassed matrix glass, gives a minimum measure of how much degassing has occurred (Anderson, 1974). For example, undegassed glass inclusions might contain a few hundred or even a few thousand parts per million S, whereas matrix glass might contain only a few tens of parts per million S; the difference in concentration is attributed to degassing before eruption. Multiplying this concentration difference by the volume of magma erupted gives a minimum mass of preeruption degassing. When the amount of SO_2 released during large explosive eruptions (as measured by TOMS or OMI) is compared to that which would have been predicted from the petrologic method, the former may be several times or even several tens of times higher than the latter and, indeed, much greater than could ever have been dissolved in the volume of magma erupted. Since this gas is erupted all at once, it must be assumed to have resided within the magma, but it must have been in discrete fluid bubbles well in excess of saturation in the melt. (Using the petrologic method on products of magmas that had discrete preeruption volatile bubbles at depth, one often observes just a small difference between S in glass inclusions and S in the matrix glass, implying minimal preeruption degassing. In fact, though, a great deal of degassing from the melt already occurred at depth – into the discrete bubble phase. As soon as CO_2-rich bubbles begin to form at depth, other volatiles including H_2O and SO_2 move preferentially into the volatile phase rather than remain dissolved in the melt.)

4.14.2.4 Viscosity, Density, and Volatile Content

Melt viscosity varies dramatically with its composition and with cooling and degassing as magma rises from a crustal reservoir and approaches the surface. Higher-silica melts show greater polymerization of silica tetrahedra, so for similar temperatures, a rhyolite might be 3–7 orders of magnitude more viscous than an alkali basalt (**Figure 5**) (Shaw, 1972; Spera, 2000 and others). At constant partial pressure of water, P_{H_2O}, the viscosity of those same melts will increase by 0.5–2 orders of magnitude per 100 °C decrease in temperature (faster in more silicic melts). Viscosity will increase by 2–4 orders of magnitude as water content drops from 8 to <1 wt.%, with changes slightly faster at lower temperatures than at higher temperature and especially fast in the last 1 wt.% of degassing. These are profound changes.

Magma viscosity (i.e., viscosity of the melt–crystal–bubble mix) also increases by orders of magnitude as crystal contents rise from 0% to 60%, following an empirical relation (in lavas) of

$$\mu_{magma} = \mu_{melt}(1 - cX)^{-2.5}$$

where $c \sim 1.67$ and X = volume fraction of phenocrysts (Marsh, 1981). However, if the growing crystals are highly elongate, with length at least >4 times width, the increase in viscosity will be slowed by shear thinning (Ishibashi, 2009).

Magma viscosity can either increase or decrease with bubbles, depending on the 'capillary number' (the ratio of viscous and surface tension stresses in a bubble). Small capillary numbers inhibit stretching of bubbles so an increase in bubbles increases magma viscosity. Large capillary numbers allow bubbles to stretch so an increase in bubbles can lower magma viscosity (Manga and Loewenberg, 2001). In either case, change is usually by less than a factor of 5 (Pal, 2003).

Magma density is a minor function of melt composition (silicic, e.g., rhyolitic magmas are slightly less dense than mafic, e.g., basaltic magmas; 2.4 vs. 2.6 g cm^{-3}). Supercritical exsolved volatile phases modestly lower magma density. A CO_2 bubble at its critical point has a density of 0.47 g cm^{-3}, while H_2O at its critical point has a density of 0.32 g cm^{-3}, but densities of both species, and mixtures thereof, are ~1 g cm^{-3} at most magmatic temperatures and pressures. A much stronger influence comes from volatiles that are already in a gas state, in magma at shallow depth at pressures less than a few tens of megapascal. Recalling the ideal gas law,

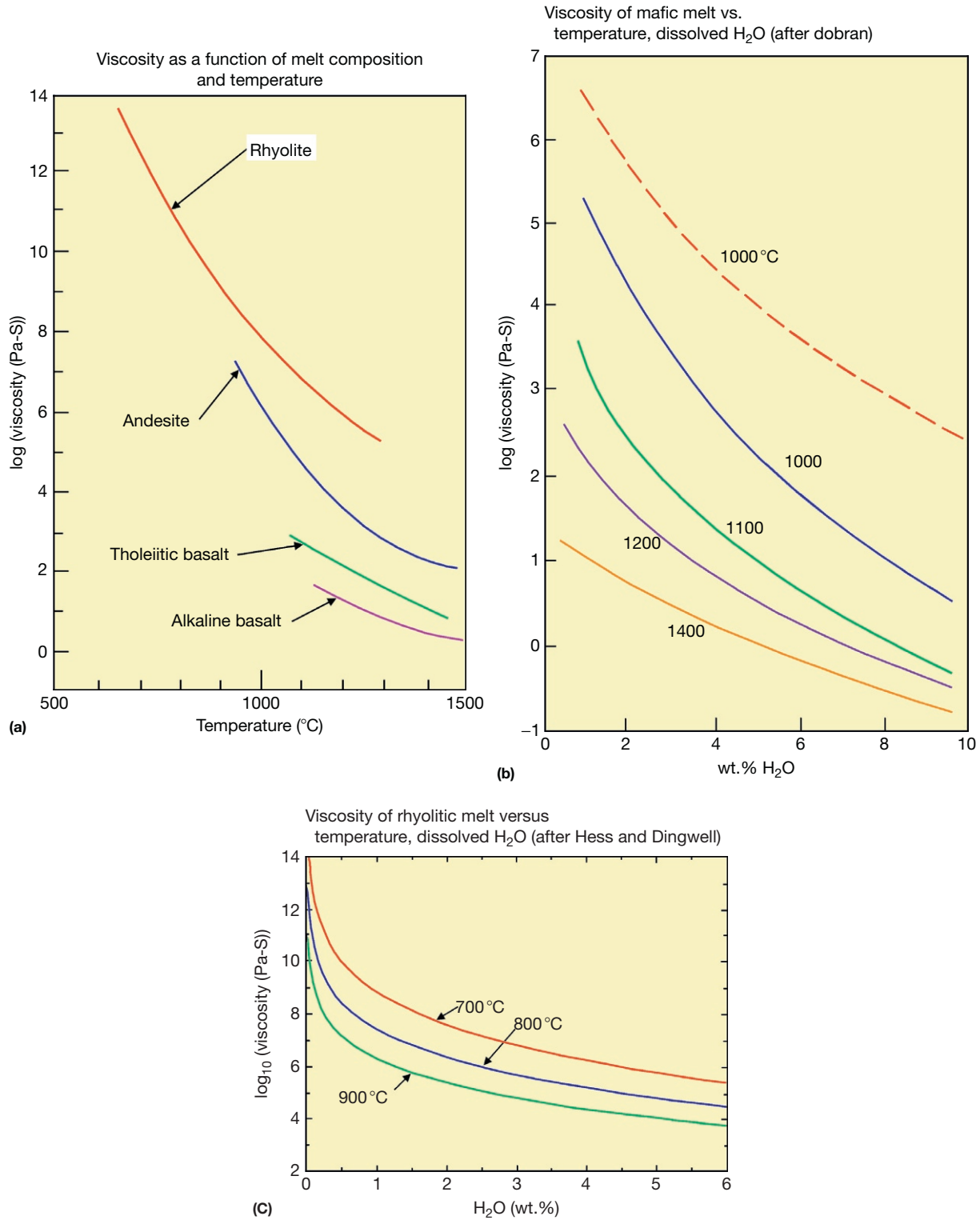


Figure 5 Viscosity of magma varies by orders of magnitude depending on SiO_2 and H_2O content and can increase by orders of magnitude as it rises to the surface, losing H_2O and growing crystals. (a) Viscosity versus SiO_2 content and temperature; (b,c) viscosity versus H_2O content and temperature, for mafic and rhyolitic melts. Reproduced from Wallace P and Anderson Jr AT (2000) Volatiles in magmas. In: Sigurdsson H, Houghton BF, McNutt SR, Rymer H, and Stix J (eds.) *Encyclopedia of Volcanoes*, pp. 149–170. San Diego, CA: Academic Press; Dobran F (2001) *Volcanic Processes: Mechanisms in Material Transport*, 590 pp. New York: Kluwer/Academic/Plenum; and Hess K-U and Dingwell DB (1996) Viscosities of hydrous leucogranitic melts: A non-Arrhenian model. *American Mineralogist* 81: 1297–1300.

$PV = nRT$, pressure and volume are inversely related when temperature and number of gas molecules are constant. If depth and confining pressure decrease, bubbles will expand and the density of magma can decrease dramatically. We can visualize this with the same bottle of soda pop that is considered earlier. The density of soda pop with only dissolved gases is close to 1.0 g cm^{-3}. Once it foams, the density can drop to a small fraction of that.

Volatile contents play into both viscosity and density. Exsolution of water from melt has the effect of increasing viscosity of the remaining melt. Because degassing of melt also induces microlite crystallization, another major contributor to increased viscosity, degassing is especially effective in sharply increasing 'magma' viscosity. To some extent, this is a self-limiting process because the viscosity of the degassed outer carapace of magma will increase so much that it acts to block further degassing, but that carapace is subject to cracking. New bubbles, if they remain in the magma, will increase or decrease viscosity slightly and have a major effect on its density.

4.14.3 Intrusions and Convection

4.14.3.1 Orientation, Geometry, and Dimensions of Intrusions

The shapes and orientations of magma conduits are functions of principal tectonic stresses and preexisting structures. In cases of intrusion into a homogeneous medium within a tectonic stress field, intrusions will be as roughly vertical planar bodies ('dikes') and perpendicular to the minimum principal stress direction σ_3. As intrusions reach closer to the surface, a vertical axis often becomes the minimum principal external stress direction, and there is a tendency for intrusions to become vertical cylinders. However, in regions of extensional stress, σ_3 remains horizontal and intrusions may persist as dikes all the way to the surface. Where magma has limited buoyancy or encounters an impenetrable horizontal layer during its ascent, it can spread laterally as broadly horizontal planar sheet ('sill').

Basaltic dikes range from a few tens of centimeters to several tens of meters in width, most less than a few meters. Composite dikes (dikes consisting of several intrusions, each one wedging into the previous one) can be wider, as can 'dike swarms' (sets of many subparallel dikes that have intruded the same area over time, each one solidifying before the next). Silicic dikes can be several meters to several tens of meters in width, rarely up to ~100 m wide. Cylindrical intrusions of magma termed plugs, often silicic, can be tens to hundreds of meters in diameter. Sills can vary from several meters to several tens (rarely, hundreds) of meters in thickness. Like composite dikes, sills may inflate by repeated, closely spaced intrusions.

4.14.3.2 Drivers of One-Way Magma Ascent

Conventional wisdom is that magma follows a one-way path upward to eruption. Forces that drive magma upward are (1) buoyancy, (2) lithostatic pressure, and (3) magmatic overpressure. Buoyancy is effective only in fluid–fluid interaction where downward-directed return flow can occur. Lithostatic pressure constantly squeezes the sides of a magma intrusion, and that magma will try to move up to a zone of lower lithostatic pressure. The tops of many crustal magma reservoirs are at 3–6 km, where lithostatic pressures are between 100 and 200 MPa; speculative reasons for this fact are raised in Section 4.14.9. Magmatic overpressure is internal pressure in excess of lithostatic pressure, caused by volatile expansion at shallow depths or by the addition of magma from greater depth. Like lithostatic pressure, it acts in all directions and not just upward, but the easiest relief of this pressure is usually upward.

In a few cases, lithostatic pressure may be augmented by compressive tectonic strain (literally, squeezing magma upward from cracks or from ductile rock at depth). Compressive strain of about 0.1 MPa beneath Mount Pinatubo as a result of the 1990 Luzon earthquake might have squeezed basaltic melt upward – either from the lower crust or, perhaps more likely, throughout the whole length of a deep magma feeder system – until it intruded a residual dacitic reservoir (Bautista et al., 1996; Newhall et al., 2002).

4.14.3.3 Drivers of Convective Magma Ascent with Return Flow

The conventional wisdom of one-way transport is hard to reconcile with the so-called excess degassing (Andres et al., 1991), a common occurrence in which volcanoes emit far more SO_2 and presumably other volcanic gases than could have been dissolved in modest volumes of magma that are erupted over the same period. A combination of field observations and modeling (Carbone and Poland, 2012; Kazahaya et al., 1994, 2002, 2004; Shinohara, 2008; Stevenson and Blake, 1998; Witham, 2011; Witter et al., 2004, 2005) suggests that in many, perhaps most of these cases, the extra gas came from magma that is convecting within the conduit. Low confining pressure allows the top of a magma column to develop into a permeable foam, probably just a few hundred meters below the surface. Gas escapes and the residual degassed, collapsed foam sinks back down through the conduit (Figure 6). Gravitational settling of these blebs of degassed magma is what drives convection, by displacing gas-rich magma upward (Bergantz and Ni, 1999). This is not a 'conveyor belt' convection cell with continuity between upwelling and downwelling limbs, where higher viscosities of the degassed magma would stop the convection. Instead, it is a modified, 'lava-lamp' style of convection in which the viscosity of the downgoing blebs is irrelevant. Sinking blebs can even be solid and the process will still operate.

Magma foams are known from their 'frozen' equivalents, including very low-density, fluffy 'reticulite' pumice. Modeling by Witter (2003), Molina et al. (2012), and others indicates that mafic magmas can and probably will convect in conduits 1–5 m diameter, and Witter found that relatively hot, volatile-rich silicic magmas will do the same in conduits 5–20 m in diameter. Relatively, high alkali content of silicic magmas will also reduce viscosity and increase the chance for convection. Rates of SO_2 gas emission constrain rates of upward flow during routine, continuous convection to be in the order of millimeters per second to centimeters per second. This has important implications for both seismic and deformation precursors of eruptions. Where magma is convecting freely in a conduit, any seismic and deformation precursors to eruption will be subtle and occur only shortly before eruption. Little or

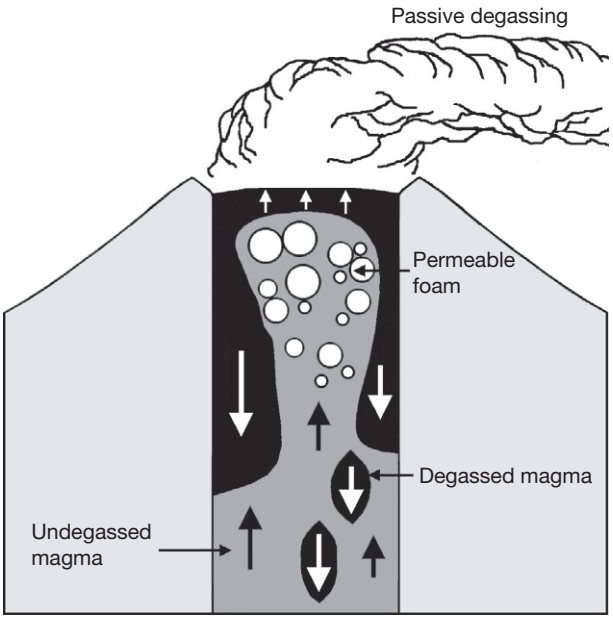


Figure 6 'Lava-lamp' convection in a conduit. Gas-rich magma develops a head of foam near the surface. Gas escapes, and the degassing foam collapses. Now dense, it sinks back down through the column, displacing gas-rich magma upward. Once the process begins, it will be driven primarily by sinking blebs, unless or until the supply of fresh gas-rich magma into the conduit becomes too slow to sustain the process. Adapted from Kazahaya K, Shinohara H, and Saito G (2002) Degassing process of Satsuma–Iwojima volcano, Japan: Supply of volatile components from a deep magma chamber. *Earth Planets Space* 54: 327–335.

no rock needs to be fractured before magma can erupt. Empirically, this is observed at frequently active, relatively open-vent volcanoes, often with mafic to intermediate magmas. Steady-state gas supply to the near surface, balanced perfectly by gas leakage (degassing), can be tipped in favor of gas supply by a fresh intrusion of magma, gas, and heat at depth and a corresponding increase in the rate of convection. This imbalance, causing gas pressures to build slightly, will produce very subtle seismic and deformation precursors to eruptions – typically recordable only at or near a crater rim just minutes before an explosive eruption, for example, at Sakurajima (Iguchi et al., 2008; Kamo et al., 1994), Suwanosejima (Iguchi et al., 2008), or Mount Merapi (Voight et al., 2000). The same subtlety applies at all open conduits, even where magma ascent appears to be one way upward, for example, at Soufriere Hills, Montserrat (Voight et al., 1999). Other shallow processes that may also contribute to near-surface, late-stage overpressure are given by Burgisser et al. (2011).

In contrast, seismic and deformation precursors where magma in the conduit is neither convecting nor otherwise open can be expected to be more pronounced, with significant magma pressurization and rock breakage as magma breaks a new path to the surface.

Aside from observed degassing, is there other evidence for convection of magma in conduits? Might there be seismic and/or electromagnetic evidence of magma moving through the 'lava-lamp' convection 'cell.' Has this been found? If not, it probably will be. Do we find large phenocrysts that show oscillatory compositional zonation to indicate that they rode a convection elevator up and down the conduit multiple times? Perhaps, as at Mount Erebus, Antarctica (Dunbar et al., 1994), though most other cases of oscillatory zoning are equivocal and might alternatively be explained by fluctuations in magma water content or temperature without convection. We might expect that U–Th ages of plagioclase crystals in convecting magma might span a wide range of ages, and indeed such ages at Mount St. Helens range from 10^3 years to just a few years (Cooper and Reid, 2003), but neither gas emission nor other evidence points to convection in the conduit of Mount St. Helens. Perhaps, the strongest evidence will be found in degassed groundmass glass. In products of the 2000 eruption of Mayon Volcano, Arpa et al. (2006) documented abundant blebs of degassed andesite with microcrystalline groundmass wrapped within darker, less degassed lava of the same composition. Degassing was inferred from textures and confirmed by direct analysis of the glasses (Cabato et al., 2012). It would be interesting to reexamine more and less crystalline groundmasses of Mount Merapi lavas described by Hammer et al. (2000), particularly in light of the findings of Allard et al. (1995) about excess degassing, which suggests convection of magma in the conduit.

4.14.3.4 Magma Mixing: Conditions, Evidence, and Effects

Mafic magmas arriving from the mantle and lower crust often encounter more silicic magmas en route to the surface. This phenomenon is especially common at arc and continental volcanoes at which the long-term supply rates of mafic magma are sufficiently high (10^6–10^8 m^3 $year^{-1}$) to generate and sustain silicic magmas by partial melting of crust and by fractional crystallization and other differentiation processes in long-lived magma reservoirs (Shaw, 1985). Such systems have a thermal memory of the previous intrusions, and many have large silicic magma bodies (<5–1000 s of km^3) at depths 5–15 km beneath the surface.

In order for magmas to mingle or mix, their viscosities must be within approximately one order of magnitude of each other (Sparks and Marshall, 1986). A number of models of mingling and mixing have been proposed. One mechanism involves underplating of basaltic magma beneath a silicic reservoir, with thermally driven convection in one or both layers and mixing along a subhorizontal interface (Sparks and Marshall, 1986; Sparks et al., 1977). Another mechanism involves intrusion of mafic magma into silicic magma. Rising mafic magma may have sufficient momentum to simply intrude right into overlying silicic magma (Bergantz and Breidenthal, 2001). In a third mechanism, cooling and partial crystallization of mafic magma when it encounters less dense silicic magma will induce vesiculation of the mafic magma and make it less dense than the silicic magma, and buoyancy will let mafic magma 'tunnel' vertically into the silicic body. Mingling will occur along a subvertical interface (Bergantz and Breidenthal, 2001; Eichelberger, 1980). A fourth mechanism, mixing induced by squeezing of two rising magmas within a conduit, was suggested by Koyaguchi and Blake (1989).

Mingling and mixing of two magmas is clearly evident in the hybrid products. Blebs of the intruding, vesiculated basalt may be quenched and preserved as inclusions in the

hybrid magma, with distinctive open-lattice ('diktytaxitic') textures (Eichelberger, 1980; Pallister et al., 1996). Further shear can cause quenched blebs to mechanically disintegrate. Incomplete mixing ('mingling') produces a marble-cake swirl of the two components. If mafic melt mixes completely with silicic melt, compositionally intermediate melts develop. Both mingled and mixed magmas show disequilibrium juxtaposition of otherwise incompatible minerals, for example, olivine and quartz, and reaction rims on minerals from the higher-temperature end-member or resorption of the rims of plagioclase from the lower-temperature end-member.

Magma mixing is not known to have a direct geophysical signal. However, changes of temperature and oxygen fugacity associated with mixing can substantially change the solubilities of gases in both magmas. If solubilities are decreased enough, bubbles will form and one should expect seismic and deformation signals associated with buoyant rise of magma and expansion or pressurization of bubbles as that magma rises and lithostatic pressure decreases. One of the best-documented instances of magma mixing began at Mount Pinatubo several months before its eruption in June 1991 (Pallister et al., 1996; M. Coombs, personal communication, 2002); the latest mixing occurred <1 week before eruption (Rutherford and Devine, 1996). Water-rich basalt intruded (tunneled into) and mixed with sluggish dacite magma, early in a chain of events that, on 15 June 1991, led to eruption of 5 km^3 of dacitic magma, the second largest eruption in the twentieth century.

4.14.3.5 Geospeedometers of Magma Ascent, Convection, and Mixing

Despite what we might expect from a dike propagating toward the surface, there are very few known cases in which earthquake hypocenters have migrated upward. A broad ~3 km upward migration of hypocenters may have occurred at Mount Pinatubo (Philippines) in 6 days in June 1991 (~20 m h^{-1}) (Harlow et al., 1996). That migration was apparent in fast playback of hypocenters in VolQuake (Hoblitt et al., 1996a) though less obvious in relocated hypocenters (Rowe et al., 2001). In 1975, at Plosky Tolbachik (Kamchatka), Fedotov et al. (1980) traced seismicity (and, by inference, basaltic dike intrusion) from 20 km to the surface at an average rate of about 0.1 km h^{-1}. In both cases, hypocenters spanned a significant part of the conduit length at any given time but moved broadly upward. Other reported cases are Kilauea east rift basalt at 0.1–1.5 km h^{-1} (Klein et al., 1987), east of Izu basalt at 1 km h^{-1} (Hayashi and Morita, 2003), and Mount Unzen dacite at 3 m h^{-1} (Nakada et al., 1999).

In principle, the tip of a dike would also be a point source of deformation (Mogi, 1958) and this, too, could be tracked upward. There are a few instances in which tiltmeters or strainmeters have recorded traces consistent with upward migration (e.g., Linde et al., 1993). Sequential radar interferograms can in principle show upward migration of a deformation source but more commonly show inflation and deflation of a stationary source, as at Mount Westdahl (Lu et al., 2003).

A different geospeedometer – applicable during and perhaps before eruptions – considers the diameter of dense ultramafic or mafic xenoliths that get erupted, calculates their likely settling rate, and infers that rate as a minimum upward velocity of magma in order to bring these xenoliths to the surface (Spera, 1984). In the 1801 eruption of Hualalai, Hawaii, magma ascent was apparently >30 m h^{-1}.

Several petrologic geospeedometers exist, all relying on the fact that minerals that are taken out of equilibrium will develop diffusion profiles near crystal rims, or reaction rims (e.g., hornblende will develop pyroxene and Fe–Ti oxide rims), and the travel time for that magma from reservoir (equilibrium) to the surface (quenching) is modeled as the time required for the observed diffusion profile or thickness of reaction rim to grow in theory and/or in lab experiments (Coombs et al., 2000; Costa et al., 2008; Rutherford and Gardner, 2000). A nice recent example shows magma recharge at Santorini caldera from <100 years to only a few months before the giant Minoan eruption (Druitt et al., 2012). Martin et al. (2008) used similar methods to infer recharge ~1 month before the much smaller 1925 eruption of the modern Kameni dome complex that has growth subsequently within the same caldera. Sometimes, the absence of zoning or reaction rims can also be used to judge the maximum ascent time on the premise that zoning or rims would have developed given more time. If the maximum ascent time is combined with erupted volume, one can estimate a minimum discharge rate and thus minimum diameter for the conduit.

In the discussion of convection, earlier, we noted that rates of ascent can be judged from gas flux rates and assumptions about conduit diameter and the concentration of gas in the magma. These estimates suggest rates of meters to tens of meter per hour (Witter, 2003).

By similar arguments, a known volume rate of extrusion and an assumed conduit diameter will give an ascent rate. At Mount St. Helens in the 1980s, mass eruption rates during explosive events implied ascent rates of 1–2 m s^{-1} (3600–7200 m h^{-1}). During nonexplosive dome extrusion in the 1980s, and again in 2004–2006, ascent rates were tens to hundreds of meter per hour (Pallister et al., 2008; Sherrod et al., 2008).

A final way to estimate ascent rate – during eruptions – is through microlite number density (MND), if one assumes equilibrium vesiculation and steady-state flow and that microlite density is proportional to water exsolution rate (Toramaru et al., 2008). MNDs in products of Izu-Oshima (1986) and Mount Unzen (1991–95) were consistent with independent estimates of ascent rate during those eruptions. Toramaru (2006) also tried to use bubble number density (BND) to measure ascent rate during eruptions but found that the number densities were too high for any plausible magma ascent rate. He concluded that BND measures instead the rate of rapid decompression and downward propagation of a vesiculation front. The method presumes volatile undersaturation until decompression begins.

Time from magma mixing to eruption can be estimated by the thickness of reaction rims or by the preservation versus loss of disequilibrium mineral pairs that were juxtaposed by mixing. From Arenal Volcano, olivine reaction rims of at least 11-year growth are described by Coombs and Gardner (2004), and, at the opposite end of the time spectrum, the preservation of disequilibrium Fe–Ti oxides at Mount St. Helens, Mount Pinatubo, Mount Unzen, and Soufrière Hills or cummingtonite at Mount Pinatubo indicates maximum

times from mixing to eruption of only a few days (Devine et al., 1997, 2003; Pallister et al., 1996, 2008; Rutherford and Devine, 1996). In another approach, Blake and Fink (2000) proposed a conceptual framework from which time frames of mixing might eventually be estimated from the relative sizes and degrees of deformation of mafic enclaves that get 'frozen' upon being mixed into cooler silicic magma.

4.14.4 Pressurization Within and Around Magma

4.14.4.1 Increase in Overpressure

Overpressure, internal pressure in excess of lithostatic or magmastatic pressure, will be increased by the following.

4.14.4.1.1 Volatile accumulation

Volatile accumulation is to eruptive potential as accumulated strain is to tectonic earthquakes. To the extent that there are 'eruption cycles,' they are driven by patterns of net volatile accumulation. Recall that under the discussion of gas budgets, it is postulated that magma and gas supply to most volcanoes is relatively frequent and in small increments. At some volcanoes, it may be almost continuous at all timescales; at others, it may be intermittent over timescales of days to years but constant over timescales of decades. In either case, the interval of magma supply from depth, into a shallow crustal reservoir, is probably quite a bit shorter than the eruption interval.

Plots of cumulative volume of magma erupted versus time for various volcanoes can show time predictability, volume predictability, or no predictability, that is, completely Poisson behavior (Koyama and Yoshida, 1994). Most volcanoes that have open vents and leak gas and magma fairly freely are time-predictable, as the threshold of internal pressure that needs to be exceeded is low and constant. Volcanoes whose conduits freeze up between eruptions are often volume predictable. At such volcanoes, explosivity correlates with volume erupted, both being controlled by the accumulated mass of volatiles that have not leaked out before magma reaches close to the surface. Most of the largest explosive eruptions – 'Plinian eruptions' – are of magma that not only is saturated with volatiles but also has accumulated 10–100x more volatiles (by weight) in bubble form (e.g., Mount Pinatubo; Gerlach et al., 1996; Wallace, 2001). This is the basis for the oft-repeated dictum that a volcano that has been dormant for hundreds of years is likely to produce a large explosive eruption. This dictum is true only if the conduit has been closed and gas has been accumulating, not leaking. Mount Pinatubo was just such a volcano and accumulated most of the incoming gas for 500 years. Iwojima in Japan has had almost as long a repose period and plenty of magma supply but shows no signs of erupting magma – because it is exceptionally leaky (Newhall, 2004; Newhall et al., 1994, 2003; Ukawa et al., 2005).

If volatiles accumulate in excess of saturation, forming bubble phases, further accumulation will gradually increase pressure on the walls of the container. If at typical magma reservoir depths, exsolved volatiles are supercritical and the pressure increase derives simply because of the increase in volume and the relative incompressibility of a supercritical fluid. If accumulation exceeds saturation close to the surface, volatiles will be subcritical and hence in gas form. Overpressure increases when volatile bubbles form, but the volume of the container is held constant by the country rock or plug. The increase in pressure will push some volatiles back into solution, but surface tension of small bubbles prevents full resolution. Volatiles in gas phase are much more compressible than those existing as supercritical fluids, but they are not infinitely compressible, hence, even they can cause pressure buildup.

4.14.4.1.2 Magma ascent and volatile expansion

As magma ascends and lithostatic pressure decreases, overpressure of the gas increases and only the strength of viscous magma and country rock will hold it in check. The overpressure of this contained gas will drive magma to intrude through country rock. Eventually, if the stress applied by the internal gas pressure exceeds the strength of the container, the container will fail and magma will erupt.

4.14.4.1.3 Heating and/or compression of groundwater

Groundwater in hydrothermal systems and even more distant can be confined, that is, unable to move freely away from any pressure gradient. Confinement is often of sedimentary origin in which sediment of a future aquifer is sandwiched between layers of less permeable sediment. Faulting can also confine aquifers, and within hydrothermal systems, the deposition of silica and iron oxide in pore space can create 'self-seals,' trapping the underlying hydrothermal waters (Fournier, 1999, 2007).

Confined groundwater, hot or cold, can add significantly to pressures in and around a magma body. If heated, it expands and increases pore water pressures (Reid, 2004), though slow diffusion of heat through groundwater will keep effects quite local or slow. Thermally induced hydraulic pressurization may be greatest at the tip of a rising body of magma – that is, a wedge driving open cracks.

If porous host rock undergoes sudden compression due to nearby magma intrusion, its pore water pressures will jump quickly (Matsumoto et al., 2002; Newhall et al., 2001; Roeloffs and Linde, 2007; Sato et al., 2001; Shibata and Akita, 2001). Observations of research wells at Mount Usu in Japan and Krafla volcano in Iceland have shown up to 100 m immediate rise in water level, corresponding to a notable 1 MPa increase in pore water pressure. This pressure bleeds off in a matter of days but can briefly trigger seismicity for 5–20 km around a volcano (White and Power, 2001; R. White, personal communication, 2012) and/or sharply decrease slope stability as probably occurred at Mount St. Helens in 1980, Mount Unzen in 1792, and many other volcanoes that have massive sector collapses (Day, 1996; Voight and Elsworth, 1997).

4.14.4.1.4 Lithostatic pressure

As described earlier, lithostatic pressure is the main source of pressure within magma bodies at depth. Lithostatic pressure squeezes the sides of a magma body at the same time it 'weighs heavily' on that magma. Squeezing of this body will be greater at its base than at its top, hence creating a driving force for upward intrusion. Furthermore, if a large body of magma feeds upward into a narrow feeder (e.g., a dike), magma in the conduit will act like a manometer (as in the straw of a slightly squeezed juice box), and stress will be concentrated at the tip of that dike that will help the magma break its way up through country rock.

4.14.4.1.5 Regional tectonics: quasistatic and dynamic strain

Any compression added by quasistatic or dynamic strain will enhance the squeezing mentioned in the preceding text. Conversely, decompression, if transient (dynamic), will briefly lower the solubility of gases and allow nucleation and expansion of bubbles in the magma (Nishimura, 2004; Shimomura et al., 2006). That effect will be ephemeral, as most volatiles will be redissolved back into the melt as soon as equilibration is reestablished; however, surface tension will preserve some small volatile bubbles and prevent complete reequilibration.

4.14.4.2 Decrease in Overpressure

Overpressure in and around magma will be diminished by the following:

4.14.4.2.1 Gas leaks

The effect of gas leaks from magma is like that of gas leaks from soda pop: the magma goes flat. For details, see earlier discussion, Giggenbach (1996) and Jaupart (1998).

4.14.4.2.2 Eruptions

Volatiles tend to accumulate in the uppermost parts of magma conduits and reservoirs and will be depleted by eruptions that, naturally, tap those parts. Eruptions are major 'gas leaks.'

4.14.4.2.3 Groundwater diffusion and dissipation of pore water pressures; Breakage of self-seals

A pressurized hydrothermal fluid surrounding magma will eventually bleed off its pressure through porous flow or fractures. If the cause of pressurization was the development of silica-rich self-seals, fracturing of those seals from within or by regional earthquakes will rapidly depressurize the source fluid but may briefly pump up pressures along connected, surrounding fractures. The concept is described by Fournier (1999, 2007).

4.14.4.3 Dynamic Balance between Pressurization and Depressurization

In summary, a variety of processes act to increase or decrease pressures within magma and surrounding hydrothermal systems. The increases and decreases may be in delicate balance This delicate balance often appears during ascent of gas-rich magma from a crustal reservoir toward the surface. Volatile exsolution occurs as solubilities in melt decrease, and both newly exsolved and any bubbles already existing in the reservoir exert increasing internal pressure as lithostatic pressure decreases. If gas is able to leak out and decrease internal pressure as fast as that pressure is increased from ascent, there will be no net increase in gas pressure and magma may arrive at the surface with little or no explosive potential or not even reach the surface (Moran et al., 2011). Conversely, if magma ascent is rapid and pressures build faster than they can bleed off, an explosive eruption will result. A positive correlation of dome extrusion rate (and inferred ascent rate) with explosivity, reported from Soufrière Hills volcano, has now been corroborated and generalized.

A parallel exists in hydrothermal systems. Confined groundwater can be pressurized by the addition of insoluble gas, by heating (from that and from soluble gas), and, most effectively, by mechanical compression of the aquifer by nearby magma intrusion. If these pressurization processes are slow, porous and fracture outflow may prevent any rise of pore water pressure. On the other hand, there are clearly some instances in which water pressures rise sharply, whether at the tip of a rising dike or more broadly, and resulting hydrofracturing may facilitate magma ascent and/or phreatic explosions.

4.14.5 Volcanic Unrest

This section is a brief overview of the nature, magnitude, and frequency of 'observable' physical and chemical changes that occur between eruptions and especially as magma intrudes toward the surface. Principal measuring techniques are also mentioned. Useful general references include UNESCO (1972), Tilling (1989), Carroll and Holloway (1994), McGuire et al. (1995), Scarpa and Tilling (1996), and Sigurdsson et al. (2000). Readers might note that gas emissions are described ahead of seismicity and ground deformation, opposite of the usual sequence. This is for two reasons: First, gas emissions are often the first potentially detectable sign of unrest. Second, increases of gas and hydrothermal pressures are what drive volcanic seismicity, ground deformation, and magma ascent.

4.14.5.1 Gas Emissions

In situ collection of gas samples from fumaroles, for subsequent analysis in a laboratory, has the advantage of collecting the whole gas and allowing the use of a range of laboratory analytic instruments (Symonds et al., 1994). However, for monitoring purposes, it has the disadvantages of collections being infrequent and often dangerous. Remote measurements from a safer distance like the foot of a volcano, from aircraft, or even from satellites have the advantage of being more frequent and safer but are usually of just one or a few gas species.

The most easily and widely measured indicator of magma degassing is sulfur dioxide (SO_2). It is much more abundant in magmatic gas than in the atmosphere and can be measured *in situ* or remotely by now-small and relatively inexpensive UV spectrometers. Early work (Stoiber and Jepsen, 1973; Stoiber et al., 1983) showed that SO_2 emission increased before eruptions as its solubility in rising magma decreased. Subsequent work has found that this pattern is sometimes complicated by path effects. One such effect is temporary blockage of SO_2 exit from magma by sharp increase in viscosity of magma and/or by quenching and solidification of a carapace on the magma. This effect is thought to explain brief decreases in SO_2 emission shortly before explosions of Galeras volcano (Fischer et al., 1996; Zapata et al., 1997), at Fuego volcano (Waite et al., 2013), and perhaps at Mount Pinatubo as well (Daag et al., 1996). Another path effect is the absorption of SO_2 into groundwater, in which it is highly soluble (Doukas and Gerlach, 1995; Symonds et al., 2001). The apparent absence of SO_2 before eruptions, once thought to indicate the absence

of magma, is now understood in some cases to reflect absorption (scrubbing) of SO_2 into groundwater. Scrubbing persists as long as the gas must pass through water-saturated country rock; it ends and SO_2 emission can rise sharply after hot gas has boiled open a dry chimney through the groundwater.

Increasingly, CO_2 gas is also monitored – directly from fumaroles, by remote sensing (FTIR), and by measurement of CO_2 concentration and flux through soil and crater lakes (e.g., Chiodini et al., 2005; Perez et al., 2011). CO_2 has the lowest solubility of any of the major gases in silicate melt, so it is the first to be released – starting when magma is as deep as ~20 km depth. It is also much less soluble in groundwater than SO_2, so at many volcanoes, it leaks out of the slopes of the volcano (diffuse degassing) and can be measured in soil. At Mount Usu, an increase in diffuse CO_2 outgassing appeared 6 months before seismic or geodetic changes were detected (Hernandez et al., 2001). At Mammoth Mountain in California, diffuse CO_2 outgassing was first noticed in late 1989 and early 1990, months after seismicity in mid-1989 signaled a magma intrusion at ~2–10 km depth, and rose again during an increase in seismicity at 10–20 km depth throughout 1997 and at 4–8 km depth in late November 1997 (Gerlach et al., 1998; Hill et al., 2003; McGee et al., 2000). The apparent delay in CO_2 in 1989–90 may simply reflect delayed sampling; there was no delay in 1997 (T. Gerlach, oral communication).

Because the solubility of SO_2 in silicate melt is greater than that of CO_2 and less than that of HCl or HF, rising magma might produce CO_2 at first, then increasing SO_2/CO_2 ratios, and eventually increasing HCl/SO_2 or HF/SO_2 ratios (e.g., Noguchi and Kamiya, 1963). This simple solubility-related progression is a nice conceptual framework for monitoring and might even offer prospects of estimating magma depth from gas emissions. However, for magmas that are already gas-saturated in their reservoirs, even relatively soluble gases that remain in melt will partition preferentially into bubble phases rather than remain in the silicate melt, and a complete suite of magmatic gases might be released together.

4.14.5.2 Seismicity

Because the author is a geologist rather than a seismologist and because **Chapter 4.15** gives readers a modern view of volcanic seismology, I originally thought to skip this section entirely. However, it might be interesting for seismologists to have a glimpse of how this one nonseismologist struggles to understand their field and to relate it to other topics in this chapter.

Most volcanic seismicity falls into categories described by Minakami (1960), Shimozuru (1972), McNutt (1996, 2000a, 2005), Zobin (2011), and **Chapter 4.15**. Volcanotectonic (VT) earthquakes result from brittle fracture (of country rock) beneath, within, and near the volcano, caused by the rise and pressurization of magma or induced by related increases in adjacent pore water pressures. Most VT events are typical, high-frequency events and are individually indistinguishable from small tectonic earthquakes. Indeed, some may be nothing more than tectonic earthquakes along faults beneath the volcano. What distinguishes VT events from strictly tectonic events are their common occurrence in swarms rather than in mainshock–aftershock sequences, their location beneath or near active volcanoes, and their common association with other indications of magma intrusion. *b*-Values of VT earthquakes tend to be higher than for tectonic events (Sanchez et al., 2004) because of heterogeneous structure and limited fault lengths and often-elevated pore pressures in hydrothermal fluids (McNutt, 2005; Sanchez et al., 2004).

Some VT seismicity that is at some distance from a vent may hint that there is magma storage out at that distance. Independent evidence from seismic or muon tomography, seismic interferometry, or deformation suggests that some magma reservoirs are out where this VT seismicity occurs, far from the vent. Suggested examples include Mount Asama (~8 km from the vent; Nagaoka et al., 2012), Shinmoe-dake (~6 km; Nakao et al., 2012 and in review), and Mount Unzen (~15 km; Nakada et al., 1999).

However, another subset of VT earthquakes that occur at distances of a few up to 20 km from their summit are far beyond the margins of any known magma body. These are called 'distal VTs' by White and Power (2001). Centered 17 km from Guagua Pichincha volcano, the north Quito swarm turned on and off as the hydrothermal system of the volcano pressurized and then erupted, respectively (Legrand et al., 2002). Near Soufrière Hills and Mount Pinatubo volcanoes, swarms 3 and 5 km from the respective summits occurred during early stages of their eruptions, as pressures in the magma reservoir and surrounding hydrothermal system increased. There is no necessary implication that magma itself is present where distal VTs are occurring nor even in hydraulic connection with the hypocentral area. The correlation can be through propagation of mechanical strain that acts on locally confined aquifers.

Decades ago, Minakami (1960) recognized a general progression from high-frequency, impulsive-onset VT (or, in his terminology, A-type) earthquakes at depths >1 km to upward to low-frequency (his B-type) earthquakes at <1 km depth that were often more emergent as well as having lower frequency content. Minakami's concept of a standard progression of earthquake types as magma rises remains useful and was nicely updated into a generic earthquake swarm model by McNutt (1996, 2000b).

Subsequent work suggests that 'B-type earthquakes' – shallow and of relatively low frequency – are actually of multiple origins. Some reflect source processes as discussed in the next paragraph, and others reflect attenuation of high frequencies and other path effects as emphasized by Minakami (1974) and Malone (1983). Ishihara and Iguchi (1989), Iguchi (1989), and Tsuruga et al. (1997) distinguished two to three subcategories of 'B-type' earthquakes at Sakurajima, representing both fracturing and subsurface explosions. Dominant frequencies vary (1–5 Hz for BL and 5–8 Hz for BH), but all seem to originate ≤3 km below the crater.

Some low-frequency earthquakes seem to occur by rupture of relatively ductile material, either hydrothermally altered rock or nearly solid, viscous magma (Gonnermann and Manga, 2003; Goto, 1999; Iverson, 2008; Moore et al., 2008; Ramos et al., 1999; Thomas and Neuberg, 2012; Tuffen and Dingwell, 2004). If strain rate is high enough, even ductile materials can rupture. A subset of these low-frequency earthquakes ('hybrids') have a high-frequency onset followed by a low-frequency tail, and these commonly occur within or just beneath growing lava domes. Rupture may begin in the brittle

carapace and propagate inward to softer material. Alternatively, hybrids might reflect occurrence of an earthquake in brittle rock that sets nearby fluid into oscillation (McNutt, 2000a) or result from complex paths rather than complex sources (Harrington and Brodsky, 2007; Kedar et al., 1996).

Another class of low-frequency earthquakes is sometimes termed a long-period (LP) earthquake to convey an added character – a relatively monochromatic character, dominantly 1–3 Hz. Some of these events are strikingly monochromatic, such as the so-called 'tornillo' earthquakes of Galeras volcano and elsewhere (Gil Cruz and Chouet, 1997; Gómez and Torres, 1997). Chouet (1992, 1996a,b) modeled these as the result of sudden jetting of pressurized gas through cracks, followed by resonance of the crack that just opened and shut. In some but not all cases, there is visual corroboration of steam jetting shortly after LP events. Initially, the occurrence of LP events in the buildup toward explosive eruptions (e.g., at Redoubt, 1989–90) led to an inference that these reflected gas pressures within the magma and explosive potential (Chouet, 1996a). Subsequent observation of many such events in the aftermath of eruptions, when conduits are relatively open and streaming gas, suggests jetting of gas through cracks but not necessarily under great pressure. Hellweg (2003) suggested an interesting alternative that the monochromatic or harmonic character of some such events might be a path effect, resulting from scattering polarization of signals.

Some low-frequency (LP) events occur deep beneath volcanoes. At Mount Pinatubo, deep LPs (DLPs) were approximately 35 km deep and were observed in late May and early June, only a week before magma erupted (White, 1996), but evidence for earlier magma mixing suggests that other LPs might have occurred before seismic monitoring was established. Beneath Kilauea, deep LPs are 30–60 km deep (Koyanagi et al., 1987), and beneath Fuji, 10–20 km deep (Nakamichi et al., 2004; Ukawa, 2005). A number of examples are noted by Power et al. (2004). DLPs are thought to reflect resupply of magma from depth into shallower reservoirs. They do not by themselves indicate imminent eruption.

With increasing use of broadband seismometers at volcanoes, there are increasing reports of very long-period (VLP) earthquakes with periods of up to several tens of seconds (e.g., Kumagai et al., 2003; Kawakatsu and Yamamoto, 2015). Some are thought to occur when slugs of fluid (magma, water, or gas) pass through constrictions. The very long period reflects the time it takes for fluid head to pass through the constriction.

The last major category of volcanic seismicity is volcanic tremor – periods of continuous vibration lasting from minutes to years (Konstantinou and Schlindwein, 2002; McNutt, 1992). The rms amplitude (~energy) of tremor, normalized for station distance, magnification, and signal spreading, is reported as reduced displacement, RD (Aki et al., 1977; Fehler, 1983). Some tremor evolves from or devolves into a succession of many small earthquakes, for example, at Mount St. Helens (Malone and Qamar, 1984) or Soufrière Hills, Montserrat (Neuberg, 2000), and is apparently just overlapping of those earthquakes.

Other tremor begins as a low-level continuous signal, grows and is sustained, and eventually fades away as it began. Some is broad-spectrum; other tremors can be of remarkably narrow spectrum with harmonics. One suggested origin, in addition to overlapping of discrete events, is nonlinear oscillation induced by flow of magma (or hydrothermal fluid) through conduits (cracks) (Balmforth et al., 2005; Julian, 1994, 2000). A second mechanism is excitation and resonance of a fluid-filled crack, irrespective of whether fluid is flowing (Chouet, 1992, 1996a,b). A third possible mechanism is magma wagging, invoked by Jellinek and Bercovici (2011) to explain tremor during explosive eruptions. Wagging refers to alternate compression and restoration of 'gas spring' force in a bubble-filled outer annulus of the conduit.

Generation of tremor by a fourth mechanism, hydrothermal boiling gets less attention than it probably deserves. Leet (1988) described experiments on 'subcooled boiling,' that is, boiling in which bubbles form where groundwater is in contact with magma or hot gas but then implode within cold surrounding water. Subcooled boiling appeared to be capable of generating strong tremor. At least some instances of strong tremor with reduced displacement $>30\ cm^2$ are associated with phreatic eruptions and with times when we know that magma is intruding and probably boiling groundwater, for example, at Mount St. Helens in April–May 1980 and during a short but strong episode on 2 October 2004. Some such tremor is also 'banded,' turning on for a few hours, off for a few hours, and back on and off for days, strongly suggestive of geyser activity (e.g., at Karkar volcano; McKee et al., 1981). Elegant confirmation of boiling and bubbles as a source of tremor comes from Kieffer (1984) and Kedar et al. (1996, 1998), who correlated pressure pulses within the Old Faithful (Yellowstone) water column and tremor recorded by seismometers at the surface.

4.14.5.3 Ground Deformation

Ground deforms around magma in two ways – elastically in response to pressure (magmatic or hydrothermal) and inelastically as it is fractured and physically shoved aside by magma or if the flank of a volcano is collapsing (e.g., Johnson, 1987). Dzurisin (2003, 2007) and Segall (2010) offer excellent reviews of methods and interpretation. Elastic deformation is a normal consequence of magma intrusion and pressurization at depth and can be aseismic. Interferometric synthetic aperture radar (InSAR) monitoring of volcanoes has shown a number of volcanoes that are being deformed by relatively deep sources (probably, resupply of basaltic magma) without accompanying seismicity, for example, at Mount Westdahl, Alaska (Dzurisin and Lu, 2007; Lu et al., 2003). Another good example is in the southwest corner of the Three Sisters volcanic complex in Oregon, where deformation began as early as 1996 and was not known to have induced any seismicity until a small, shallow swarm in March 2004 (Dzurisin et al., 2006, 2009; Wicks et al., 2002).

The shape and gradients of the deformation field indicate the depth and shape of the deformation source (Cervelli, 2000; Dieterich and Decker, 1975; Mogi, 1958; Okada, 1992; Segall, 2010). It is beyond the scope of this chapter to describe the relationship between sources and deformation fields in detail, but a few general relations can be noted. Uplift from a point source (such as the rising tip of a magma column) forms concentric contours. The half-width of uplift, defined as the diameter of the contour of half the maximum uplift, is 3/2

times the depth of the point source. Tilt resulting from a point source shows its maximum value at a radial horizontal distance 0.5 times the depth of the point source. Different patterns result from dikes. These models, developed by Mogi, Okada, and others on theoretical grounds, are remarkably robust and consistent with seismic and other evidence for magma depths and shapes. Inelastic deformation immediately around a magma intrusion is just extreme strain, exceeding the tensile strength of the rock. For example, grabens typically form above a near-surface dike and the north flank of Mount St. Helens bulged irreversibly northward as a viscous dacite cryptodome (hidden dome) intruded into its north flank.

Iwo-jima (Japan) shows an interesting combination of elastic deformation and inelastic deformation. Over the past 400 years, the floor of the submarine Iwo-jima caldera has risen more than 120 m above sea level (Kaizuka et al., 1985; Newhall et al., 2003; Oyagi and Inokuchi, 1985; Ukawa et al., 2005). Episodic magma intrusion of approximately 10^7 m^3 $year^{-1}$ is the driver. Some of the deformation is elastic, in which episodes of inflation once every decade or two are partially reversed over succeeding years by deflation of the central part of the island. Some recent inflation well away from the center of the caldera (Ukawa et al., 2005) is more likely to be inelastic, and net deformation since the time of caldera formation has been inelastic, broad uplift, and numerous faults with offsets of several meters that cut post-WWII era runways (Oyagi and Inokuchi, 1985). Iwo-jima is an actively resurging dome and is also remarkable for its efficient degassing of magma that (at least for now) seems to prevent magmatic eruption (Newhall et al., 2003). Some of the deformation clearly involves intrusion, possible pressurization, and degassing of magma; other deformation probably involves pressurization and depressurization of the surrounding hydrothermal system. Similarly rapid, episodic resurgent uplift occurs at Siwi caldera (Vanuatu) (Chen et al., 1995; Métrich et al., 2011).

Because some nonvolcanic landslides are known to show accelerating creep shortly before failure, those monitoring the bulge at Mount St. Helens in 1980 were hoping for a similar indication of imminent failure. It did not occur. If anything, creep slowed slightly in 3 days immediately before the failure (Lipman et al., 1981). Apparently, resistance along the basal plane of the landslide block(s) was overcome suddenly.

4.14.5.4 Other Changes: Gravitational, Magnetic, Electrical, Thermal, and Hydrologic

The acceleration of gravity changes with changes in elevation or change of nearby mass. The most commonly used unit, a Gal, equals 0.01 ms^{-2}, or about 10^{-3} times the acceleration of gravity on Earth. Gravity will decrease by 0.3086 mGal for every meter of increase in elevation but also increases by ~0.1 mGal due to the addition of 1 m of rock (Rymer, 1996). At volcanoes, the expected change in gravity from uplift or subsidence, without change in nearby mass, that is, the Bouguer-corrected free air gradient (BCFAG), is in the order of 0.200–0.250 mGal m^{-1} depending on the assumptions about source geometry and density of the volcanic pile (Gottsmann and Rymer, 2002; Rymer and Williams-Jones, 2000). If the gradient of gravity change with elevation is significantly higher or lower, a mass change is also implied. An increase along the BCFAG would be interpreted as subsidence without mass change; a faster increase could be from intrusion of magma that is denser than the rock into which it intrudes, constant-volume densification of magma by collapse of vesicles and coupled inflow, and/or a compression-induced rise in water table (increasing density just beneath the gravimeter). A decrease along the BCFAG would be interpreted as uplift without mass change; a faster decrease could suggest drainback of magma (rare), vesiculation of magma, or dilatation-induced drop in water table. Independent geodetic measurements and modeling of the combined deformation-gravity field are needed to discriminate between these possible contributors (Rymer and Williams-Jones, 2000). Frequent or, better yet, continuous measurements are needed to capture short-lived, transient events. With GPS and other geodetic instruments, we can have continuous measures of deformation, but only a very few volcanoes have had continuously recording gravimeters (Mount Etna, Mount Vesuvius, Mount Merapi, Mount Asama, Iwate, Miyake-jima, and Kilauea), and even fewer still have them.

The strength of the Earth's magnetic field at its surface ranges from 25 to 65 μT (0.25 to 0.65 G), strongest near the magnetic poles and weakest near the magnetic equator. Magnetic fields are known to change by as much as ten or more nanoteslas at some volcanoes and during some intrusions (Johnston, 1989; Sasai et al., 1990; Zlotnicki, 1995). Heating can slowly demagnetize rock near an intrusion. Stress associated with intrusions can induce piezomagnetic effects, in the order of 1 nT per 10 MPa of overpressure (Zlotnicki, 1995). Still other changes may be of electrokinetic origin, when a conductive fluid (hydrothermal or magma) is moving (Fitterman, 1978; Mizutani et al., 1976).

Magma intrusions can change the distribution of resistivity within a volcano, by changing porosity and permeability of a hydrothermal reservoir or conductivity of a hydrothermal fluid. Lenat (1995) gave a useful review of these changes and ways to measure them. One particularly interesting example is at Kilauea, where active-source electromagnetic monitoring saw change in phase and amplitude of signals during passage of a dike intrusion, and remeasurements of self-potential across a rift zone also showed change after a dike intrusion (Jackson, 1988; Jackson et al., 1985). Similar changes have been recorded at Piton de la Fournaise and Etna volcanoes (Lenat, 1995) and at Miyake-jima (Zlotnicki et al., 2003).

Minor shifts in the locus of thermal activity occur frequently at volcanoes, probably related to self-sealing and changing patterns of permeability within the hydrothermal system. Most of these are insignificant. Increases in ground, water, or gas temperatures by several degrees or even tens of degrees are more significant, potentially caused by increased flux of hot gas. Crater lakes are particularly useful integrators of thermal and chemical flux from hot gas streaming up into the lake (Varekamp and Rowe, 2000), so Trunk and Bernard (2008) devised a useful application of ASTER satellite imagery to monitor changes in crater lake temperatures. Fumaroles at or near the boiling temperature of water rarely change temperature because they are buffered by boiling of groundwater. High-temperature, drier fumaroles (some as hot as 950 °C) can have many-degree swings in temperature with increases or decreases in the flux of gas from magma. Large increases in volume of steaming, as at Mount Baker (1975), Mount Pinatubo (1991), Turrialba volcano (1997–present), or Garbuna Group of

volcanoes (2004), can be another indication of hot gas from a fresh magma intrusion heating water of a hydrothermal system.

Finally, changes in water level or spring discharge occur at some volcanoes, principally in response to mechanical compression or dilatation of confined aquifers (Newhall et al., 2001; Roeloffs and Linde, 2007). Most reports of change in water level are anecdotal, but good measurements have been obtained in Japan and in Long Valley Caldera. As mentioned earlier, increases in pore water pressure can induce seismicity at or near a volcano, cause modest inflation, and destabilize the flanks of volcanoes if pressurization occurs in a widespread, outward-dipping weak layer such as a layer of weathered ash (Day, 1996; Reid, 2004; Voight and Elsworth, 1997). Gravity changes at Nisyros (Gottsmann et al., 2005) and at Soufrière Hills volcano, Montserrat (Hautmann et al., 2010), were attributed to steam- and deformation-induced changes in the water table, respectively.

4.14.6 Volcanic Eruptions

Historically, volcanologists believed that nearly all eruptions were triggered internally, that is, when pressures within magma exceeded the tensile strength of country rock. No doubt this is true for many eruptions. But, especially since the 1992 Landers earthquake, evidence has built that some eruptions are triggered externally (Brodsky and Prejean, 2005; Hill et al., 2002; Hill and Prejean, 2007, this volume; Manga and Brodsky, 2006; Schmincke, 2004; Walter, 2007; Walter and Amelung, 2007). Here is a list of potential eruption triggers.

4.14.6.1 Eruption Triggering Mechanisms

4.14.6.1.1 Internal

4.14.6.1.1.1 Intrusion of fresh magma (and its gas) from depth, driven by partial melting and by buoyant forces and/or lithostatic squeezing

These drivers exert modest upward and lateral pressures that can overcome tensile strength of country rock and lead to eruption. Also, gas dissolved within that incoming magma is added to preexisting magma, potentially increasing buoyancy and internal gas pressure (overpressure) to the point of failure. As noted earlier, overpressure also forces gas bubbles back into solution, but surface tension of small bubbles prevents return to full equilibrium.

4.14.6.1.1.2 Buildup of gas pressure as magma rises

The solubility of gases decreases with decreasing lithostatic pressure, and if gas has no room to expand, its overpressure will tend to rise. The tendency toward pressure increase with magma ascent holds true regardless of whether magma is on one-way ascent or in the upwelling limb of two-way convection. This and the preceding mechanism **(4.14.6.1.1.1)** are probably common triggers of eruptions.

4.14.6.1.1.3 In situ buildup of gas pressure from crystallization, as in 'second boiling'

Crystallization concentrates volatile phases in the residual, interstitial silicate melt. When those concentrations rise above gas solubility, bubbles will form and increase in either volume or pressure. This process, called 'second boiling,' may trigger relatively small eruptions.

4.14.6.1.1.4 Mixing of fresh magma with preexisting magma

Mixing may add gas to the preexisting magma and can also change the oxygen fugacity (f_O). Sulfur gases in particular have a solubility minimum in temperature–f_O space, so if mixing pushes magma toward this minimum in gas solubility, more gas will exsolve (Kress, 1997). Addition of fresh mafic magma usually raises the temperature of the preexisting magma, in turn decreasing the solubility of most gas in that preexisting magma. Also, thermally or gas-induced density changes in the preexisting magma cause gas-bearing magma from the base of the reservoir to rise and exsolve still more gas. Any combination of these effects can trigger an eruption.

4.14.6.1.2 External

4.14.6.1.2.1 Regional earthquakes

Following the 1992 Landers earthquake and a number of more recent earthquakes, local seismicity and strain increased at Long Valley Caldera and other large calderas. Possible mechanisms for such response are discussed in Hill et al. (2002), Manga and Brodsky (2006), Hill and Prejean (2007, this volume), and Bebbington and Marzocchi (2011). Many involve geothermal waters.

Although most regional earthquakes do not trigger eruptions, nor are most eruptions triggered by earthquakes, there are several dozen known cases in which a strong earthquake was followed 'shortly' by an eruption, for example, eruptions of Cordón Caulle (Puyehue) in Chile within hours of major earthquakes in 1921 and 1960 (Katsui and Katz, 1967; Lara et al., 2004) and again 15 months after a strong earthquake in 2010. 'Shortly' is relative to the normal repose periods (interevent times) of the volcano, hours versus decades in the case of Cordón Caulle. In 1707, the large and most recent eruption of Fuji began just 49 days after the M8.4 Hoei earthquake. In 1917, San Salvador volcano began steaming immediately and erupted for several months after an M6.5 earthquake occurred just west of the capitol city. Another apparent case of earthquake-induced eruption is that of Mount Pinatubo, Philippines, which became restless within weeks and erupted months after a nearby M7.8 earthquake, after (as best as is known) being quiet for the preceding 500 years.

How might regional earthquakes trigger eruptions? Two familiar analogies suggest one possibility. If one shakes a container of soda pop or adds Mentos and caps it, many new bubbles will nucleate and exert overpressure. Shaking the soda pop creates local decompressed patches in which bubbles can nucleate; Mentos also stimulates bubble nucleation. Because the increased internal pressure will tend to force volatiles back into solution, any lasting increase in overpressure depends on the presence of small bubbles, surface tension, and rigid reservoir walls (Nishimura, 2004; Shimomura et al., 2006). Shaking-induced bubble nucleation in magma was suggested as a possible eruption trigger by Manga and Brodsky (2006), as were shaking-induced rise and pressurization of bubbles in hydrothermal water by Linde et al. (1994).

Alternatively, there might be local faulting that relieves confining pressure. This has been inferred in Chile, where

eruptions of Cordón Caulle have almost certainly been triggered by large regional earthquakes (Bonali et al., 2012; Lara et al., 2004). In Kamchatka, at Karymsky, and the adjacent Akademii Nauk caldera in 1996, magma accumulation may have stressed a regional fault, which, upon rupture, unclamped fractures and allowed magma to rise and erupt 2 days later (Walter, 2007). Where eruptions occur in regions of volumetric dilatation, unclamping of faults and/or decompression and exsolution of CO_2 may bring magma closer to eruption (Walter and Amelung, 2007).

In regions of volumetric compression from a nearby earthquake, magma might be squeezed upward, as was inferred at Mount Pinatubo (Bautista et al., 1996) In that case, the M7.8 Luzon earthquake on 16 July 1990 caused +0.1 MPa static compression throughout the crust beneath Mount Pinatubo. Based on petrologic evidence, fresh basalt intruded into the preexisting, subvolcanic dacite reservoir 6 months after the earthquake and possibly before then as well. The basalt might have risen as a diapir (discrete batch) all the way from the lower crust, or it might have been squeezed and coalesced from interconnected magma-filled cracks throughout the thickness of the crust. Deep LP earthquakes (35 km deep), thought to be related to basalt movement at depth, are known to have occurred by late May 1991 and may also have occurred earlier. A mixture of the basalt and dacite first erupted on 7 June 1991.

Finally, eruptions that are delayed by weeks or months may be induced if a dense mat of crystals is dislodged from the roof or walls of a body of magma, sinks, and displaces an equal volume of gas-rich magma upward (Bergantz and Ni, 1999; Hill et al., 2002; Manga and Brodsky, 2006). Interstitial melt of a crystal mush might be squeezed upward, just as in the action of a French press for coffee.

In summary, a relatively small percentage of eruptions are triggered by regional earthquakes, but under such varied scenarios of shaking or local stress change, various triggering mechanisms seem likely.

4.14.6.1.2.2 Unloading, rapidly as by a massive landslide off a volcano's flank, or slowly, as by glacial retreat

Decrease of confining pressure induces exsolution of volatiles and the related chain of increased buoyancy and/or increased internal pressure. A classic example of an eruption induced by a large landslide is that of Mount St. Helens in 1980. A statistical increase in eruptions following glacial retreat has been postulated in Iceland, but some caution is needed to distinguish between real postglacial increases in volcanism and better postglacial preservation of deposits (Maclennan et al., 2002; Van Vliet-Lanoë et al., 2005) (see also Jellinek et al., 2004; Watt et al., 2013).

Mastin (1994) made a related proposal: that rainfall-induced fracturing of hot rock can induce phreatic and eventually magmatic eruptions, working from the top down.

4.14.6.1.2.3 Earth tides

Some volcanoes show a slight statistical preference to erupt on or around fortnightly tidal maxima (or, in a few cases, fortnightly minima), and a few show a preference to erupt on or near semidiurnal tidal maxima (or minima) (Dzurisin, 1980; Emter, 1997; McNutt and Beavan, 1981, 1987). Most volcanoes whose eruptions can be triggered by Earth tides have relatively open conduits, relatively low-viscosity magma, and frequent, small eruptions, for example, Kilauea, Pavlof Volcano, and Mayon Volcano. The exact mechanisms by which Earth tides trigger eruptions are not known, but possibilities include slightly enhanced exsolution during semidiurnal tidal decompression and/or a slightly increased rate of magma ascent. Hydrofracturing might also be involved, as seismicity in areas of high stress and high hydrothermal pore pressures can also correlate with Earth tides (Glasby and Kasahara, 2001).

Statistical preference to erupt on tidal maxima (or minima) does not indicate a significant increase of eruption probability at each tidal cycle. Volcanoes that are not ready to erupt will not be affected by tidal cycles. Only those volcanoes that are ready to erupt, and that require only the slightest additional push, might show an influence from the tides.

4.14.6.2 Explosive versus Effusive Eruptions: Fragmentation of Magma

When magma finally nears the surface, its gas content and viscosity will exert important controls over the style of eruption. An overly simplistic view is that magma with high gas content will explode, producing ash and larger pyroclasts, while magma with low gas content will erupt gently (effusively), producing lava flows. A simple analogy is the initial opening of soda pop or champagne, charged with gas and explosive if opened suddenly. Once open, it starts to 'go flat' and, soon, can produce only a gentle overflow. Viscosity in its simplest role limits how easily gas can escape from magma, that is, how easily it can go flat. While gas is certainly a necessary component of explosive eruptions and while viscous magma tends to erupt more explosively than fluid magma, neither all gas-rich magmas nor all viscous magmas will explode.

As magma is rising up a conduit, the concentration of volatiles is controlled by its initial gas concentration and en route leakage, the latter being controlled in large part by viscosity, the ability of bubbles to separate from melt, and ascent time. Wall-rock permeability may also be a factor. At Soufrière Hills volcano, faster ascent brought gas-rich but viscous magma to the near surface faster than it could degas, so explosive eruptions ensued (Sparks and Young, 2002).

More generally, the explosivity of an eruption is an expression of the amount and rates of exsolution and bubble growth (Sparks, 1978, 2003). When magma reaches within several tens to hundreds of meters of the surface, the dominant controls on rates of exsolution and bubble growth become

- the remaining volatile content, in melt and as discrete bubbles;
- any sudden rupture and release of confining pressure;
- bubble nucleation rate;
- melt viscosity, which limits
  - how fast dissolved volatiles can diffuse through melt into bubbles (water in silicic (viscous) to mafic (nonviscous) melts diffuses from 10^{-1} to 10^{3} mm^2 s^{-1}) (Behrens et al., 2004; Zhang and Behrens, 2000; Zhang et al., 2003),
  - how fast bubbles can expand,

- whether melt as it is stretched by bubble expansion will fragment or continue to stretch, and
- permeability of any foam that develops.

As magma expands into foam, either gas can bleed off and leave the remaining magma intact, or it can expand and stretch the melt walls until they fail (Cashman et al., 2000). Low-viscosity bubble walls can stretch and thin to a point of breakage; higher viscosity and rapid shear will cause bubble walls to fragment by brittle rupture (Ichihara, 2008). In laboratory experiments with rapid decompression (intentionally rupturing a membrane in a pressure vessel), bubbles expand rapidly and typically induce fragmentation (Alidibirov and Dingwell, 1996; Ichihara et al., 2002; Kameda et al., 2008; Martel et al., 2000; Phillips et al., 1995). A typical vol.% of bubbles before fragmentation is 70%, reachable by expansion of just 0.1 wt.% gas from 100 MPa (3 km) to 0.1 MPa (surface), and far exceeded with higher wt.%'s of gas. If this vol.% of bubbles is exceeded rapidly, the stressed and stretched molten glass walls of those bubbles break quickly and the magma fragments (blows itself to pieces). However, if the decompression rate is slower, shearing-induced permeability may prevent explosions (Cashman and Mangan, 1994; Eichelberger, 1995; Eichelberger et al., 1986; Klug and Cashman, 1996; Mueller et al., 2008). Burgisser and Gardner (2005) estimated that decompression equivalent to an ascent rate of magma of less than 0.35 m s^{-1} could prevent explosions.

Nucleation and diffusion rates are a minor issue if most of the gas is already in separate bubble phase (in extreme cases, magma can have 100 times more volatiles in pre-ascent bubbles than dissolved in the melt). Where there is not already a separate bubble phase, the presence or absence of other impurities like crystals is important. With impurities as bubble nucleation sites, exsolution can begin immediately upon saturation, so there is a long, slow, weakly explosive process of 'heterogeneous' bubble nucleation and growth. Without impurities, magma becomes oversaturated with gas and then bubbles nucleate 'homogeneously,' suddenly, and explosively (Mangan and Sisson, 2000; Mourtada-Bonnefoi and Laporte, 2004).

Thus, the explosivity and eruption – reflecting the rates of bubble nucleation, growth, and magma fragmentation – is a complex function of

- initial gas content and melt and magma viscosity;
- the presence or absence of tiny microlites to cause heterogeneous bubble nucleation;
- the presence or absence of a preexisting discrete bubble phase and its number density; and
- the effective permeability of the magma, magma foam, and wall rock and hence the efficiency with which magma can bleed off excess gas pressure shortly before and during eruptions.

4.14.6.3 Measures of Magnitude and Intensity (Volume Dense Rock Equivalent, Mass Discharge Rate, and VEI)

The concepts of magnitude ('bigness') and intensity of eruptions are different than magnitude and intensity of earthquakes. 'Magnitude' refers broadly to the volume or mass of products erupted, and 'intensity' is the rate at which products are erupted. Volume is usually expressed in terms of cubic meters or cubic kilometers, with a qualifier to indicate whether that is the volume of deposit ('bulk volume,' thickness times areal extent, including bubble and intergranular space) or the volume of magma that erupted ('DRE,' or dense rock equivalent, corrected to a bubble-free or prevesiculation volume of magma). Tsuya (1955) proposed the first volume-based magnitude scale for eruptions and applied it to the sum of lavas and pyroclastic deposits.

Conversion from bulk volume back to DRE volume is made by multiplying bulk volume by the ratio of average density of deposit to the density of bubble-free magma. Typically, lavas have densities close to that of magma, and pyroclastic deposits have densities around half that of magma.

Mass of magma erupted, as required for the Pyle (2000) magnitude scale of explosive eruptions, is estimated by a similar process, multiplying the bulk volume as measured in the field by the average bulk density of deposits.

Mass discharge rate is typically expressed in mass erupted/unit time, for example, kg s^{-1}. It is estimated by multiplying erupted volume times density and dividing by the duration of the eruption. A theoretical and empirical relation has been demonstrated in explosive eruptions between column height and mass discharge rate (Carey and Sigurdsson, 1989; Wilson, 1976; Wilson et al., 1980), so the instantaneous intensity of an eruption can also be estimated from column height. Bursik (2001) and Woodhouse et al. (2013) reminded us that strong wind will shear and reduce the height of a column. Another empirical relation, between reduced displacement of eruption tremor and column height, was found to give another estimate of column height during eruptions (McNutt, 2000b). Prejean and Brodsky (2011) added a physical basis for this relation by successfully working backward from column heights of Kasatochi and Augustine eruptions to expected tremor amplitudes.

To estimate volume and mass discharge rate requires mapping of deposits and/or detailed visual, satellite, or seismic observations during an eruption. This is possible for many modern eruptions but difficult to impossible for many older eruptions. To describe the explosive magnitude of older eruptions semiquantitatively, Newhall and Self (1982) proposed the VEI that is based (in decreasing priority) on bulk volume of pyroclastic deposit, maximum column height, duration of eruption, and qualitative adjectives. Over all but the smallest eruptions, VEI rises one step with each order of magnitude increase in volume of deposit. As with earthquake and many other magnitude–frequency relations, there is a power-law decrease in frequency with increasing VEI (Siebert et al., 2010; Simkin, 1993; Simkin and Siebert, 2000) (**Figure 7**). Not all eruptions that occur after a long repose (e.g., after hundreds to thousands of years) are of high VEI, but nearly all high VEI eruptions follow long reposes. Not counting short intervals between the start and climax of large explosive eruptions, it appears that a long repose is a necessary but not sufficient condition for a large explosive eruption, probably because it allows enough time for gas to accumulate in excess of saturation if the volcano is not leaky.

However handy VEI is as a shorthand among volcanologists, it is not a quantitative parameter, so bulk volume, mass discharge rate, and other details should be reported directly when known.

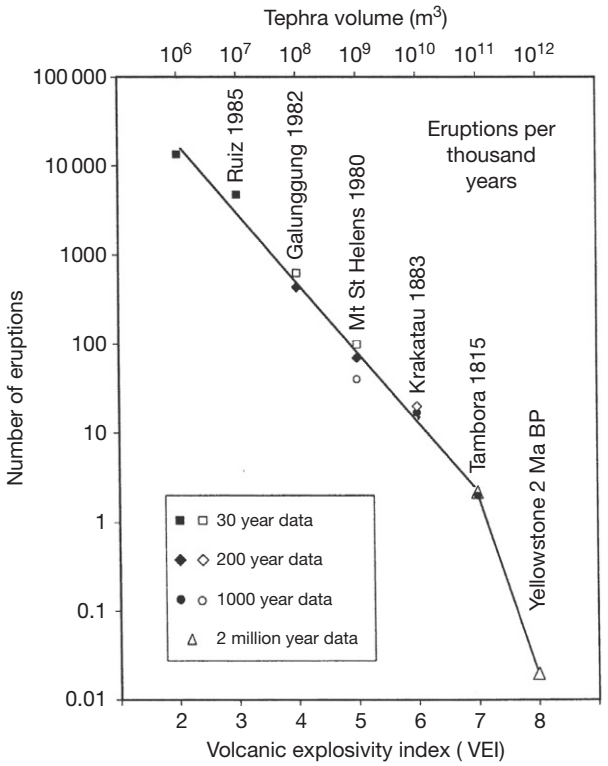


Figure 7 VEI magnitude–frequency plot. Over the range from VEI 2–8, each successively higher VEI is about six times less frequent than the preceding lower VEI. A similar magnitude–frequency relation for earthquakes is well known. Reproduced from Simkin T (1993) Terrestrial volcanism in space and time. *Annual Reviews of Earth and Planetary Sciences* 21: 427–452, updated in Simkin T and Siebert L (2000) Earth's volcanoes and eruptions: An overview. In: Sigurdsson H, Houghton BF, McNutt SR, Rymer H, and Stix J (eds.) *Encyclopedia of Volcanoes*, pp. 249–261. San Diego, CA: Academic Press.

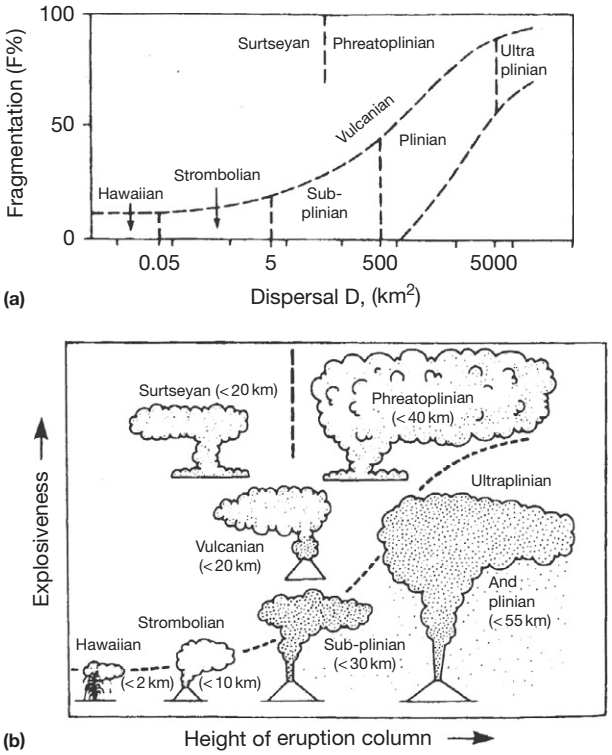


Figure 8 (a) Types of volcanic eruptions plotted in F–D space. F = fragmentation = wt.% of tephra fall that is <1 mm, along dispersal axis where $t = 0.1T_{max}$; D = dispersal = area enclosed by $0.01T_{max}$ isopach line. T = thickness of tephra fall deposit (from Cas and Wright, 1987, after Walker, 1973; Wright et al., 1980). (b) Types of volcanic eruptions as functions of column height (itself, a proxy for mass discharge rate) and explosivity (as indicated by the degree of fragmentation) (from Cas and Wright, 1987). Plot (a) is strictly descriptive; plot (b) is interpretive, though based on measurable proxies.

Walker (1973) proposed a more quantitative classification of explosive eruptions, based on two characteristics of their fall deposits:

(1) Fragmentation (F) = wt.% of pyroclasts <1 mm along dispersal axis, where thickness = 0.1 maximum thickness of deposit ($0.1T_{max}$).
(2) Dispersal (D) = area enclosed by ($0.01T_{max}$) isopach line of the tephra fall deposit.

Walker's fields of various eruption types on plots of (F) versus (D), as modified by Cas and Wright (1987) and reproduced here in **Figure 8**, correspond roughly to terminology that had already developed over preceding decades (Plinian, Strombolian, etc.) but define those eruption types more precisely.

Rutherford and Gardner (2000) and Cashman (2004, 2006) discussed the controlling effects of magma ascent rate, vesiculation, and degassing on explosivity. Ascent rates of 0.5 m s^{-1} or higher inhibit degassing en route to the surface and result in higher explosivity; those of 0.05 m s^{-1} or lower encourage en route degassing and result in relatively low explosivity. Burgisser and Gardner (2005) reached a similar conclusion: that magma ascent of <0.35 m s^{-1} could inhibit explosions. Cashman (2006) suggested correlation of magma supply rate to column height and dispersal (D) and proportion of erupted magma that is erupted explosively ('% tephra') to explosivity and fragmentation (F).

Further discussion of various measures of eruption 'bigness' may be found in Walker (1980), Carey and Sigurdsson (1989), and Pyle (2000).

4.14.6.4 Eruptive Phenomena

4.14.6.4.1 Lava flows and lava domes

When magma pours or oozes from the ground, it is renamed 'lava.' Given a slope and a viscosity low enough to flow, lava will form lava flows. Most lava flows are elongate ribbons of lava with thicknesses of 10^0–10^1 m, widths of 10^1–10^2 m, and lengths of 10^2–10^4 m. Flows an order of magnitude larger are not uncommon. Flood basalts such as the Columbia River Basalt or the Deccan Traps are erupted from long dikes and form broad, thick sheets up to 10^6 m long and wide. Basaltic and other low-silica, low-viscosity flows can travel considerable distances; more silicic and thus more viscous lava might travel only a few kilometers from its vent.

A number of different terms are used to describe the surface features and flow details of lavas. Two of the most common are

Hawaiian terms, '*aa*' and '*pahoehoe*.' *Aa* lava has a blocky, rough surface, while *pahoehoe* has a smoother, ropy surface. Various factors contribute to the difference, but the main factors are viscosity and shear rate. Higher viscosities and higher shear rates favor *aa*.

Extremely viscous lava – too viscous to flow more than a few tens or hundreds of meters from the vent – mounds up and forms 'lava domes.' Growth by extrusion of one or a succession of viscous flows is called exogenous; growth by inflation of an existing dome is called endogenous. An endogenous dome that grows just beneath and dramatically lifts the ground surface – soil, trees, and all – is called a cryptodome ('hidden dome').

Nearly all lava flows travel slowly enough that people can walk out of harm's way. Buildings and other immovable objects are generally destroyed, either by fire or by mechanical bulldozing.

Thick, extensive flows of obsidian (volcanic glass) bear special mention. None have been observed closely during formation, but they pose an interesting paradox. Glass forms by rapid quenching of melt upon eruption – usually, within seconds or minutes of eruption. The glassy character of obsidian implies rapid quenching. But many obsidian flows are so thick that they would have been expected to cool slowly, crystallizing rather than being quenched. Also, most obsidian flows are highly silicic and would have been expected to be more viscous than their broad dimensions suggest. Eichelberger et al. (1986) suggested a provocative explanation: that magma was erupted as a permeable foam, like a good foam on a beer. As gas escaped, the foam collapsed into a quenched, dense glass that spread just a short distance farther. This might give the deceptive appearance of obsidian having flowed as a solid, viscous lava over the entire distance from its vent, whereas it might instead have been a much lower-viscosity foam over most of that travel distance. A related explanation has been offered as well – that hot rhyolite magma can ascend unusually fast, retain volatiles and heat, have very little crystallization, and thus have a viscosity low enough to flow freely (Castro and Dingwell, 2009; Pallister et al., 2010) – but the paradox of rapid quenching of a thick flow remains.

4.14.6.4.2 Tephra fall

When an eruption is explosive, magma is blown out as fragments ranging in size from fine dust or sand ('ash') to large blocks or volcanic bombs. These fragments are called 'pyroclasts' (from the Greek for fire and broken) or 'tephra' (from the Greek for volcanic ash). Large fragments behave as ballistics, following parabolic trajectories and landing within a few hundred meters (rarely, up to a few kilometers) from the vent. Smaller fragments can be carried farther aloft by vigorous thermal convection. Eruption clouds from the largest explosive eruptions convect up to 50 km above the vent; small and moderate-sized eruption clouds (from VEI 1–3 eruptions) rise 0.1–20 km above the vent.

At altitude, ash or tephra-bearing clouds are typically sheared by prevailing winds and tephra rains out as a curtain elongate in the direction of the prevailing wind. For the most part, it rains out in order of decreasing particle size and density, following Stokes' law, but, if moist or electrically charged, the finest grain sizes clot (aggregate) together and fall out earlier than they otherwise would.

Light tephra fall (up to a few centimeters thick) does not directly threaten the lives of people on the ground but can cause serious damage to computers and other electronics, electrical distribution grids, heating–ventilation–air conditioning (HVAC) systems, storm drains, and vulnerable crops. Tephra fall in excess of 10 cm, especially if wet, is the threshold beyond which structural damage and roof collapse are possible. Ash clouds pose a special hazard to modern jet aircraft, whose engines run at temperatures high enough to melt ingested ash. In the cooler parts of the turbine, molten ash glazes over fuel and air nozzles, engines shut down, and planes become gliders. Serious encounters, including near crashes of jumbo jets, have led to a new, global 24/7 (around-the-clock) network of Volcanic Ash Advisory Centers (Casadevall, 1994; NOAA Satellite and Information Service, 2013; Romero, 2004). Although data are sparse, guidance from the International Civil Aviation Organization (ICAO) since 2012 is to avoid visible ash with a presumptive concentration of $\sim$2 mg m^{-3} or higher.

4.14.6.4.3 Pyroclastic flows and surges

Pyroclastic flows are avalanches of hot fragments and gas down the slopes of volcanoes, typically at 20–50 m s^{-1}. The terminology of pyroclastic flows is complicated, reflecting multiple origins of flows (Wright et al., 1980). Most pyroclastic flows form by collapse of an eruption column, collapse of a dense slug of debris erupted just a few hundred meters above a vent, or collapse of the toe of a lava flow or dome growing on a steep slope. Pyroclastic flows have a relatively dense basal avalanche that exhibits laminar flow and an overriding, expanding turbulent ash and steam cloud. Ash from the basal avalanche is elutriated just like the dust cloud from a normal rock avalanche and is almost as hot as the basal avalanche. Basal avalanches tend to follow valleys but can climb over some ridges; the overriding ash clouds are less constrained by topography. Run-out distances are typically 1–20 km but in extreme cases can reach >100 km. Because of their high speed and high temperature, pyroclastic flows are generally lethal. The worst tragedy from a pyroclastic flow occurred in 1902 on the island of Martinique, where nearly all 29 000 residents of St. Pierre died in one short-lived but deadly flow.

Pyroclastic surges are relatively dilute cousins of pyroclastic flows and indeed very similar to the upper, turbulent facies described in the preceding text. The blast that leveled forest at Mount St. Helens was, in most senses, a pyroclastic surge. Increasingly, the term pyroclastic density current (PDC) is being applied to the spectrum of events from pyroclastic flows to surges.

4.14.7 Construction and Erosion of Volcanoes

In general, eruptions build volcanoes and gravity and water conspire to erode them. Fluid magma erupts mainly as lavas that build broad, gently sloping shield volcanoes like Mauna Loa and Kilauea in Hawaii. Viscous, silicic magma forms lava domes if it degasses before eruption and explosive products (tephra fall and pyroclastic flows) if it reaches the surface still

rich in gas. Large-volume pyroclastic flows can form broad, gently sloping aprons or sheets around a source vent. Eruptions of more than 5 or 10 km^3 of magma of any composition can induce collapse into just-vacated reservoir space and form a large depression (caldera) at the surface. Magmas of intermediate composition and viscosity often produce both lavas and pyroclastic deposits, and stacks of these form steep-sided stratocones.

Construction of volcanoes, though, is offset in substantial measure by erosion. In the early stages of a volcano's lifetime, construction will outpace erosion, but eventually erosion will dominate. At intermediate stages, the two can alternate. Especially in tropical settings, erosion can match rates of construction so cones remain at roughly constant volume for many millennia. The less consolidated or indurated a deposit, the more easily it will be eroded. At Mount Pinatubo, 200 m deep preeruption canyons were filled within a day with hot, unconsolidated pumiceous deposit and then reexcavated over the course of just 2–3 years. The higher and steeper a volcano grows, or the more altered or weathered it becomes, the more it is prone to giant landsliding or 'sector' collapse, as has occurred repeatedly at Mount St. Helens and many other volcanoes (see **Section 4.14.7.2**, in the succeeding text).

4.14.7.1 Hydrothermal Alteration

Every volcano, to one degree or another, has a hydrothermal system – a system of groundwater heated by magma and circulating convectively through fractures and porous rock. Hydrothermal fluid that is heated directly by volcanic gases is acidic and reacts with country rock to form clays and other alteration minerals plus a more neutral hydrothermal fluid. Largely hidden from view, hydrothermal alteration softens and weakens a volcanic edifice and makes it prone to large-scale collapse. Some volcanoes, for example, Yellowstone, Uzon, and Mount Rainier have unusually extensive hydrothermal alteration. At Mount Rainier, hydrothermal alteration weakens the steep slopes and has been partly responsible for some giant landslides (flank collapse) that evolved downslope into clay-rich debris flows.

4.14.7.2 Sector Collapses

Two distinctive geomorphic features – horseshoe-shaped craters and fields of small hills ('hummocks') at their feet – were recognized at a handful of volcanoes before 1980. Some were thought to be the result of large, possibly slow-moving landslides; others, including small hills north of Mount Shasta, were not understood at all. Immediately upon the giant landslide of Mount St Helens in 1980, hundreds of such pairs were found and understood at once to have formed in the same, catastrophic way (Siebert, 1996; Siebert et al., 1987; Ui, 1983). Subsequent work at Mount St. Helens confirmed at least three previous collapses, two from the north flank about 3 ka BP and another from the south flank 20 ka BP (Hausback, 2000).

Although the first sector collapses were noticed from steep-sided arc stratovolcanoes, workers in Hawaii soon realized that even gently sloping shield volcanoes have also collapsed, in landslides from hundreds and even up to 5000 km^3 (Moore et al., 1994). Just one 10 min long event – the landslide of Mount St. Helens – changed the paradigm of thinking of volcanoes as fundamentally stable with rare exceptions to fundamentally unstable edifices that usually fail at least once in their lifetime.

A variety of triggers have been suggested for sector collapses (Siebert, 1996; Siebert et al., 1987). Some collapses, like Bezymianny in 1956 and Mount St. Helens in 1980, were induced by off-center cryptodome emplacement. Other volcanoes may be wedged into collapse by dike intrusions or weakened into collapse by hydrothermal alteration. Earthquakes have been suggested but not confirmed. A theme that runs through a number of collapses, and may in fact be the most effective and common trigger, is intrusion-induced increase of pore water pressure (Day, 1996). As a magma intrusion rises high into a volcanic cone, the compressional deformation field can raise pore pressures in confined aquifers by as much as 1000 kPa or more, substantially reducing normal slope stability (Newhall et al., 2001; Voight and Elsworth, 1997). Intrusion-induced, days-long increases in pore water pressure are likely to have occurred at many if not most volcanoes that have collapsed.

4.14.7.3 Concentrated Sediment–Water Flows (Lahars)

Concentrated flows of volcanic sediment and water, called by the Indonesian term 'lahars,' form by the following:

- Eruption through crater lake (e.g., Kelut, Indonesia), followed by scouring of downstream deposits,
- Rapid melting of snow and ice by pyroclastic flows and incorporation of sediment from the pyroclastic flows and also from scoured downstream deposits (e.g., Nevado del Ruiz, Colombia),
- Sector collapse of water-saturated cone, followed by either downstream evolution of avalanche into debris flow (e.g., Mount Ontake, Japan; Mount Rainier, the United States), or halting and compaction dewatering of the debris avalanche deposit (e.g., Mount St. Helens, the United States),
- Overtopping and rapid scouring failure of natural dams (e.g., Ruapehu, New Zealand; Mount Pinatubo, Philippines), and
- Heavy rainfall on fresh pyroclastic deposits (e.g., Mount Pinatubo, Philippines) Fine ash on the Earth's surface reduces infiltration rates by an order of magnitude, and watersheds become prone to sediment-laden flash floods, just as in arid climates or following forest fires.

The first two originate only during eruptions. The latter three can be syneruptive or posteruptive.

Lahars behave in ways quite unlike normal streamflow. Bulking up from clear water streamflow to >20 vol.% sediment creates hyperconcentrated flows, and further bulking up to >60 vol.% sediment creates debris flows, both termed lahars. Hyperconcentrated flow is dilute enough that it has no appreciable yield strength, yet it transports considerable volumes of sediment in short periods (Pierson and Costa, 1987). Hyperconcentrated flow is turbulent and deposition from that flow is progressive (grain by grain) though fast enough that classical stream deposit features such as cross-bedding are absent. In contrast, a debris flow exhibits

significant yield strength and laminar flow, the latter helping it to flow fast and far. Debris flows are analogous to fast-moving slurries of concrete, and the combination of their density and speed allows them to plow right through (or lift and float) buildings, bridges, and anything else in the way. Deposition from debris flows can range from progressive to sudden, en masse deposition, and resulting deposits are massive and without any internal stratification.

Lahars generate relatively high-frequency tremor, and dense seismic networks can detect and track lahars. Inexpensive 10–300 Hz geophones installed on riverbanks (LaHusen, 1996, 2005) were used with good success to detect and quantify lahars at Mount Pinatubo (Marcial et al., 1996; Tungol and Regalado, 1996) and are presently deployed at a number of volcanoes including Mount Rainier and Ruapehu to warn of approaching lahars. Short-period and broadband seismometers can also be used for this purpose (e.g., Cole et al., 2009; Doyle et al., 2011).

4.14.7.4 Normal Streamflow, Sediment Issues, and Flooding

Lahars can transport very large amounts of sediment in the years following a large explosive eruption. In the case of Mount Pinatubo, >3 km^3 was eroded from the roughly 5–6 km^3 of 1991 pyroclastic debris on the slopes of the volcano (Tungol, 2002). Most of this sediment was transported as lahars in the first 5 posteruptive years, but by 1997, 25% of declining annual sediment transport was by normal streamflow (Hayes et al., 2002). Except for one large lake-breakout lahar in 2002 and a few rain-induced lahars, most sediment transport since 2000 has been by normal streamflow. High sediment yields were also observed in the decades following the 1980 eruption of Mount St. Helens (Major, 2004), where torrential rains are rare and most sediment was carried by normal streamflow.

Copious loads of sediment are deposited in channels on the surrounding lowlands, sharply reducing their capacity to carry even modest normal runoff. Lowland flooding is a serious by-product of lahars and other heavy sediment transports after an explosive eruption.

4.14.8 People and Volcanoes

Millions of people choose to live around tropical and temperate volcanoes. Reasons include fertile soils, healthy cooler climates above tropical lowlands, and aesthetic and recreation appeal. The result is a dangerous juxtaposition of large populations with dangerous volcanic phenomena and an extra challenge for those who must forecast eruptions and avoid false alarms (Ewert and Harpel, 2004).

4.14.8.1 Risk Mitigation

The principal tools for long-range risk mitigation are hazard and risk maps (the latter integrating hazard, vulnerability, and exposure), land-use management, public education and dialog about volcanic risks, and early development and testing of emergency plans. The principal tools for immediate risk mitigation are volcano monitoring, short-range eruption forecasts, evacuations, aircraft reroutings, and other emergency measures such as installing ash filters and covering storm drains.

Can people, buildings, and infrastructure be protected against ashfall? In general, yes, though at some cost. Can flow hazards be mitigated by engineering measures? The answer is a qualified 'sometimes.' If topography is favorable, lava flows can be slowed or diverted away from critical structures. If topography is favorable and considerable time and money are invested in advance, lahars can be slowed and sediment can be sequestered by dams and related structures. Even bunker-like shelters from pyroclastic flows have been tried though with little success. However, the success rate of such efforts to control volcanic flows is low and costs are high. Usually, it makes more sense simply to avoid building in the path of volcanic flows and to move out of the way if threatened.

4.14.8.2 Eruption Forecasts

Some volcanologists distinguish between forecasts (relatively general, often probabilistic statements based on past history) and predictions (more specific deterministic statements taking special account of monitoring and interpretation of current process) (e.g., Swanson et al., 1983). In contrast, others, including the author, see a spectrum of statements that are all probabilistic to some degree and are thus 'forecasts,' based on monitoring, past history, and interpretation of current process (e.g., Marzocchi et al., 2007; Newhall, 2000).

4.14.8.2.1 Long-range forecasts

Forecasts for eruptive activity within coming months to decades are typically based on eruptive history as seen in the geologic record and in historical documents. The past is the key to the future. In a few cases, there might be robust enough clustering or other temporal patterns to infer something other than a Poisson distribution, but these are rare. From geologic and human history, we can outline the types, frequency, and probable extent of future eruptions. Such forecasts are heavily probabilistic. Hazard maps are one form of long-range forecasts, as are estimates of the probability that a particular site (e.g., a critical facility) will be hit by eruptive events within a specified time. Long-range forecasts are used for land-use planning, emergency response planning, facility siting, and other planning decisions.

4.14.8.2.2 Short-range forecasts

Forecasts for eruptive activity within minutes to months are based on a combination of geophysical, geochemical, and visual monitoring plus background information from the long-range forecast. Specific patterns of seismicity, ground deformation, gas emission, and other changes can indicate specific steps along the way toward eruption. In some cases, progress toward eruption gives a monotonic or exponential increase in monitored parameters; in other cases, values of a parameter might suddenly decrease, for example, seismic quiescence shortly before some eruptions. Patterns vary from volcano to volcano and, in some cases, from eruption to eruption, and the best short-range forecasts will be those that weigh empirical experience and real-time interpretation of processes together.

When a volcano that is preparing to erupt is known to have exhibited a wide range of explosive scales in the past, can the explosive magnitude (VEI) of its impending eruption be forecast? At present, there is no tried and true method for forecasting VEI. Indeed, precursors of the giant 1991 eruption of Mount Pinatubo were indistinguishable from precursors to much smaller eruptions until well into the eruption, just 24 h before its climax (Harlow et al., 1996; Newhall et al., 2002; Power et al., 2001).

One approach to estimating explosive 'potential' (i.e., maximum possible VEI) of an impending eruption is to calculate a gas budget since the last time gas-rich magma was reamed out of the volcano (e.g., since the time of the previous VEI ≥ 3 eruption). If during the latest repose magma has degassed apace with inferred fresh volatile supply, the author's unpublished estimates suggest that the maximum VEI will be 4 and most probably less, but if volatiles have accumulated into a discrete volatile phase, the maximum potential VEI could be 5 or higher. This approach is being tested as new eruptions permit, but it has not yet been validated. Within the bounds of maximum possible VEI, the actual VEI will depend greatly on magma ascent rate so any indication of rapid magma ascent, including rapid escalation of unrest, will favor higher rather than lower VEI.

Short-range forecasts can be of narrative form (what is expected, when, where, etc.), in the form of alert levels (e.g., color-coded alert levels; Gardner and Guffanti, 2006), or in the form of forecast windows that might be weeks long in an early or intermediate stage of unrest and days to hours long if seismicity and other parameters point to an imminent eruption (Swanson et al., 1983).

One way to integrate short- and long-range forecasts is through the use of event trees (Aspinall et al., 2002; Marzocchi and Bebbington, 2012; Marzocchi et al., 2004, 2010, 2012; Newhall and Hoblitt, 2002) and Bayesian belief networks (Aspinall et al., 2006; Hincks et al., 2006). The thought process to estimate probabilities at each node of an event tree, and make Bayesian updates of these with each new piece of information, can be very helpful in guiding scientific discussion and further work.

4.14.8.2.3 Successful forecasts and false alarms

We have already seen that volcanoes give signals that magma is rising or otherwise preparing to erupt. Some particularly gratifying successes, such as at Mount Pinatubo in 1991 and Merapi 2010, have raised expectations. On the other hand, some eruptions still occur without being forecast, from volcanoes that are not monitored at all or at which the precursors were too subtle for current monitoring. Sometimes, warning signs are detected but underestimated or otherwise misinterpreted.

An ironic challenge for volcanologists is an overabundance of precursors and how to distinguish those that will lead to eruption from those that will not. Signs that magma is rising toward the surface are similar regardless of eventual outcome. At an early stage of unrest, we may see unmistakable signs of magma rising but not known yet whether it will stall or eventually erupt, and intrusion may continue until practically the 'last minute' before stalling, as occurred at La Grande Soufrière, Guadeloupe, in 1976 (Feuillard et al., 1983; Fiske, 1984). Empirically, magma that is rising rapidly, with corresponding accelerating unrest, is more likely to erupt. Nishimura (2006) modeled accelerating ground deformation due to magma ascent without degassing, whereas degassing magma ascent produced constant-rate deformation. There may be other differences between successful and failed eruptions, also related to magma ascent rate and gas overpressures, but these are still in the research arena.

In 1999, precursors warned of an approaching eruption of Tungurahua volcano in Ecuador. The town of Baños, extremely vulnerable, was evacuated. The volcano did erupt and still erupts, but until 2006, the eruption was more spectacular than dangerous. The townspeople returned home after only a few months of evacuation and were loath to ever evacuate again. However, as the eruption continues, the magma progressively degasses and is increasing in viscosity, producing gradually more explosive eruptions and some pyroclastic flows. The lay public can see and hear the larger 'canonazo' explosions, but it is hard for them to visualize that the deposits on which their town is built are from pyroclastic flows not much older than the town! Fortunately, the scientific monitoring team has spotted precursors to larger-than-normal events, alerted the public, and saved lives in 2006 and again in 2012–2013. Notwithstanding scientists' early concern that no one would ever evacuate again, residents in the NW and W sectors have been willing to evacuate again and have lived to tell about it.

A stark illustration of another false alarm dilemma comes from Mount Vesuvius and towns around its base. The population is so large (approximately 550 000 in the 'red' danger zone) that an evacuation will require at least 1 week. Yet, it is rarely possible to be sure that a volcano will erupt until just hours to days before an eruption. The earlier the evacuation, the better the chances that it will be successful but also the greater the chances of false alarm.

Simplified 3–5 level alert schemes are now in widespread use. Their principal advantage is that civil defense and politicians can tie their mitigation plans and actions to the simple alert levels; a secondary advantage, for scientists, is that alert levels can be raised and lowered repeatedly if unrest fluctuates, without committing to a more specific forecast that might or might not prove correct. However, Winson et al. (2012) reviewed the use of alert levels across a number of countries and crises and found that in a disturbing number of cases, alert levels either were not raised until an eruption had already begun or were lowered prematurely. Most schemes are fine, but implementation has been spotty. Two other concerns about alert level schemes are that (a) lock-tight ties between alert levels and mitigation actions tend to push scientists, de facto, into making what should really be political decisions about mitigation and (b) alert levels can be misused as a crutch when we should really be trying harder to understand activity and make more specific forecasts.

4.14.8.3 Quantification and Acceptance of Risk

Quantification of risk requires quantification of hazard, value, exposure, and vulnerability, as in a formula modified from Fournier d'Albe (1979):

Risk (*probable loss*)=
Hazard (*the physical events*)
×*Value* (*population and property in the hazard zone*)
×*Vulnerability* (*% of value that would be lost if hit*)
×*Exposure* (*likelihood of value remaining at risk*, *even despite \ monitoring and warnings*)

Assessment of hazard has been discussed earlier. The Volcanic Population Index (VPI) (Ewert and Harpel, 2004) can be used as a (partial) proxy for value, and a variety of recent studies (e.g., Wilson et al., 2011, 2012) are updating and quantifying the pioneering work by Blong (1984) on the effects of eruptions on people and things.

Communities and individuals may resist evacuation and other costly mitigation measures, especially if eruptions of that volcano are infrequent or if forecasts are conspicuously uncertain. In some cases, this stems from ignorance or denial in the face of clear and present danger, but in other cases, it stems from a calculated decision to accept the risk from the volcano rather than take socially and economically costly mitigation measures. As they describe standoffs between officials who recommend evacuation and citizens who resist, media sometimes lose sight of the fact that both sides are making rational choices. The differences are simply in how much risk each side is willing to accept. Quantification of volcanic risk, even with high uncertainties, usually helps citizens and officials decide whether that risk is acceptable to them. Further discussion of this can be found in Newhall and Hoblitt (2002), Aspinall et al. (2002), Hincks et al. (2006), Marzocchi and Woo (2007), Woo (2008), and Marzocchi et al. (2012).

4.14.9 Future Directions

Brodsky (2004) offered interesting thoughts on outstanding challenges in volcanology. The author concurs and adds these additional thoughts and questions.

What causes magmas to pond and accumulate in reservoirs with tops between 3 and 6 km?

This has been termed a region of 'neutral buoyancy' by Ryan (1987). Is this the level at which decompression degassing increases viscosity sharply and causes magmas to stall? Or is 3–6 km a region in which rising magma encounters significant groundwater and is thus quenched and slowed and/or can spread by hydrofracturing of country rock? Is this a level at which lithostatic pressure that is squeezing magma upward becomes too low to force open new fractures toward the surface? Or is it a level where exsolution of volatiles forms a barrier of low-density magma that blocks ascent of following magma? Understanding when and where each of these processes is dominant may help us to understand periodicity of shallow intrusions and eruptions and to understand how long-dormant volcanoes reawaken.

Is magma supply from the mantle and lower crust continuous, and are eruptions thus controlled by buffering in shallow reservoirs and/or controlled by near-surface valves? Or, alternatively, is supply into volcanic roots episodic and, itself, a major control on when and how volcanoes erupt?

One way to address these questions is by InSAR monitoring of broad deformation, indicating resupply of magma into volcanic roots. With C-band radar, only a limited number of unvegetated high-latitude or high-altitude volcanoes could be monitored. X-band radar of TerraSAR-X is also subject to decorrelation as vegetation grows, but its 2-week return period minimizes that problem. Once L-band radar (e.g., on-board ALOS-2) is again available, we will see good data from many volcanoes, even in the tropics.

Supply of magma from depth is probably episodic if viewed to the resolution of hours to years. But as the viewing resolution blurs to decades or centuries, supply might be more or less steady. One hypothesis to be tested is that if supply is episodic and less frequent than the time required for magma in the conduit to degas, cool, and solidify after an eruption, then eruptions will occur only after fresh magma resupply and perhaps only after repeated episodes of resupply. If supply is more or less steady on the timescale of solidification of magma in conduits, then solidification will occur only when the rate of magma supply is relatively low. When supply rates are higher, conduits will stay open and filled with convecting magma, and eruptions can occur at any time or with just the slightest fresh resupply.

How, in detail, does magma degas, and how does this control eruptive style and magnitude?

This topic was introduced earlier in two contexts: leakage during magma ascent if that ascent is slow enough and leakage during the moment of eruption (modulated expansion of magma foam) if the decompression is slow enough to allow the development of shear permeability before the magma explodes. Both of these topics invite further laboratory experimentation and field confirmation. There are great opportunities for more interaction between those who measure gas emissions in the field, those who study degassing and fragmentation in the laboratory, and those who reconstruct it from the petrology of erupted products.

Specifically, field and lab investigators can test the working hypothesis that volcanoes that are plugged tight (i.e., with minimal intereruption passive degassing) can accumulate gas in excess of saturation and have high explosive potential, especially if their eventual ascent to the surface is rapid. Does all magma that feeds Plinian eruptions have an excess volatile phase? For silicic magmas, the answer may be 'yes,' but is it also 'yes' for mafic magmas? At volcanoes with open conduits, what causes the occasional explosions that are much larger than normal? Is it simply that the convection is sped up, or is homogeneous nucleation also involved?

Do early bubbles in gas-charged magma, acting as potential sites for further heterogeneous nucleation, ever reduce explosivity? In relatively viscous magmas, what is the relative importance of foaming versus shear permeability as a means to efficiently degas magma? How was strong degassing like that of the 1990s–2000s at Popocatépetl sustained even when little or no magma was erupting? How does Iwo-jima degas so efficiently, and, in particular, might groundwater and hydrofracturing play a role? What independent evidence can be found to support or reject lava-lamp style conduit convection and associated degassing at the top of an inferred 'convection cell?' Combined geologic, geophysical, and geochemical studies are needed to improve forecasts of explosive magnitude

(VEI) so that evacuations and other measures are of appropriate scale. This is a particular problem at open-vent volcanoes, for example, Tungurahua, Mayon Volcano, and Mount Merapi, where people are lulled by frequent, relatively small eruptions and can be surprised by larger ones. The same general issue of forecasting VEI arises during unrest at any volcano that exhibits a range in the VEIs of its eruptions. Can estimation of a gas budget, as in **Figure 3**, help in estimating the VEI of an impending eruption?

At what timescales are eruptions stochastic versus predictable?

Eruptions may seem to occur randomly or predictably, depending on the timescale. Many variables influence whether a volcano will erupt in a given time period (e.g., history of magma supply, gas leakage, and local stress–strain relations), so on longer timescales, eruptions may indeed be stochastic. At the same time, systematic buildup of precursors before eruptions argues for predictability. On timescales of weeks to months, volcanoes with open conduits, erupting frequently, may appear to be stochastic while those with closed conduits, erupting infrequently, appear predictable. On the other hand, if one looks closely enough at those with open vents, they, too, exhibit precursors and predictability on timescales of seconds to minutes. For forecast purposes, volcano monitoring must be scaled to the openness of the vent.

How can we best estimate critical intrinsic parameters (e.g., viscosity, water content, and overpressure) from observable parameters (seismicity, deformation, gas leaks, etc.)?

A conundrum facing volcanologists is that the parameters that directly indicate explosive potential (volatile content of magma, gas overpressures, magma viscosity) cannot be measured before eruptions. What can be measured – seismicity, deformation of the ground surface, volumetric strain, gas emissions, ground or fumarole temperatures, and a few other parameters – are at best indirect indicators of the state of the magma.

Four general approaches can help. The most direct is to estimate overpressure from deformation (many authors) and from stress axes (e.g., Miller and Savage, 2001; Roman et al., 2006).

A second approach is to track the volcano's changes (seismic, ground deformation, volumetric strain, gravity, and gas) in response to known, repeating periodic natural signals such as Earth tides. In response to tidal decompression, we might expect greater expansion response if shallow magma is charged with a separate volatile phase (high explosive potential) than if it is gas-poor. Less viscous magmas might respond more quickly than more viscous magmas. Microseismic signals and hour-to-hour gas emissions might vary with volatile concentration, state, and magma permeability.

A third approach is to use more or less constant signals, including ambient seismic noise and cosmic muon flux, to repeatedly image magma within a volcanic system. Bubbles, in particular, will significantly increase seismic attenuation and reduce velocities and magma density. Bubbles by themselves mean only magma has been able to vesiculate; bubbles combined with evidence for magma pressurization will be more worrisome.

As a fourth approach, we might try artificial sources to image magma. The main challenge here is how to focus an artificial source through relatively small target bodies of magma? Stacking of repeating artificial signals, such as in the Japanese ACROSS project (Ikuta and Yamaoka, 2004; Ikuta et al., 2002), is promising but has not been proved at volcanoes yet. Other ideas are needed!

How can we know if an impending eruption from a large volcanic system will be just a small aliquot or a colossal eruption?

Large magnitude eruptions are thought to grow from small ones by runaway vesiculation (bubble nucleation and growth), just as large earthquakes may grow from small ones by cascading rupture. Volcanic gas is concentrated in the uppermost parts of magma reservoirs. If the gas content of magma is low, an eruption might tap just a small volume from the top of a reservoir. On the other hand, if the gas content is high and the bleed rate is low, an explosive eruption may escalate. Runaway vesiculation may increase the volume and mass discharge rate of an eruption to exceptionally high values. Hoblitt et al. (1996b) interpreted the pattern of more and more closely spaced, smaller and smaller eruptions of Mount Pinatubo in 1991 leading up to the runaway climactic phase as a result of tapping magma with increasing gas concentrations, which systematically reduced the time between explosive pulses. At the same time, the vent was also getting reamed out, also decreasing resistance to eruption. Eventually, the pulse that began at 1342 h in 15 June tapped magma that was rich enough in gas to sustain and escalate the discharge. Harlow et al. (1996) reported a parallel escalation in the energy release of shallow LP earthquakes. The runaway eruption continued until the eruption tapped into deeper, gas-depleted magma and/or tapped so deep in the conduit that overlying foam and debris shut off vesiculation or the conduit was squeezed shut with the same effect.

We do not know if the patterns observed by Harlow et al. (1996) and Hoblitt et al. (1996b) are characteristic of the buildup to all large (Plinian) explosive eruptions. If they are, then we might have hours of forewarning of larger explosive events. From a practical point of view, populations must already be evacuated on the basis of any geologic evidence of past Plinian eruptions, as the confirming indications from monitoring might come too late for evacuations.

4.14.10 Conclusion

This chapter began with analogies between volcanoes and fault zones. Both are subject to stress and strain, both produce distinctive patterns of seismicity, and both exhibit 'stick–slip' behavior with similar power-law magnitude–frequency relations. In the course of this chapter, we have also seen a further analogy: that the variable degassing pattern of volcanoes, some open and freely degassing versus others plugged tight, is analogous to creeping versus locked segments of faults. Gas accumulation in magma is analogous to strain accumulation on faults, steadily accumulating unless released by passive leakage or creep. The estimation of explosive potential demands an accounting of gas accumulation versus passive degassing, over the time frame since the previous explosive eruption (typically, years to decades) and also over the time frame of current magma ascent (typically, hours to weeks).

Acknowledgment

Many thanks to reviewers of this work and its early version: Steve Malone, Emily Brodsky, and Mie Ichihara.

References

Aki K, Fehler M, and Das S (1977) Source mechanism of volcanic tremor: Fluid-driven crack models and their application to the 1963 Kilauea eruption. *Journal of Volcanology and Geothermal Research* 2: 259–287.

Alidibirov M and Dingwell DB (1996) Magma fragmentation by rapid decompression. *Nature* 380: 146–148.

Allard P (1977) Endogenous magma degassing and storage at Mount Etna. *Geophysical Research Letters* 24(17): 2219–2222.

Allard P, Carbonnelle J, Dajlevic D, Metrich N, and Sabroux J (1995) The volatile source and magma degassing budget of Merapi volcano: Evidence from high temperature gas emissions and crystal melt inclusions. In: *Merapi Decade Volcano Workshop, Yogyakarta, Indonesia*. 66 pp.

Allard P, Carbonnelle J, Metrich N, Loyer H, and Zettwoog P (1994) Sulphur output and magma degassing budget of Stromboli Volcano. *Nature* 368: 326–330.

Anderson AT Jr. (1974) The before-eruption water content of some high-alumina magmas. *Bulletin of Volcanology* 37: 530–552.

Andres RJ, Rose WI, Kyle PR, et al. (1991) Excessive sulphur dioxide emissions from Chilean volcanoes. *Journal of Volcanology and Geothermal Research* 46: 323–329.

Annen C (2009) From plutons to magma chambers: Thermal constraints on the accumulation of eruptible silicic magma in the upper crust. *Earth and Planetary Science Letters* 284: 409–416. http://dx.doi.org/10.1016/j.epsl.2009.05.006.

Arpa MCB, Bornas MAV, Abigania MIT, et al. (2006) Characterization of lava flow and pyroclasts from the February to March 2000 eruption of Mayon Volcano, Philippines. *Journal of the Geological Society of the Philippines* 61: 1–19.

Aspinall WP, Carniel R, Jaquet O, Woo G, and Hincks T (2006) Using hidden multi-state Markov models with multi-parameter volcanic data to provide empirical evidence for alert level decision-support. *Journal of Volcanology and Geothermal Research* 153: 112–124.

Aspinall WP, Loughlin SC, Michael FV, et al. (2002) The Montserrat Volcano Observatory: Its evolution, organization, role and activities. In: Druitt TH and Kokelaar BP (eds.) *The Eruption of Soufrière Hills Volcano, Montserrat from 1995 to 1999. Geological Society of London, Memoir*, vol. 21, pp. 71–91. Bath, UK: Geological Society Publishing House.

Balmforth NJ, Craster RV, and Rust AC (2005) Instability in flow through elastic conduits and volcanic tremor. *Journal of Fluid Mechanics* 527: 353–377.

Bautista BC, Bautista MLP, Stein RS, et al. (1996) Relationship of regional and local structures to Mount Pinatubo activity. In: Newhall CG and Punongbayan RS (eds.) *Fire and Mud: Eruptions and Lahars of Mount Pinatubo, Philippines*, pp. 351–370. Quezon City and Seattle: Philippine Institute of Volcanology and Seismology and University of Washington Press.

Bebbington MS and Marzocchi W (2011) Stochastic models for earthquake triggering of volcanic eruptions. *Journal of Geophysical Research* 116: B05204. http://dx.doi.org/10.1029/2010JB008114.

Behrens H, Zhang Y, and Xu Z (2004) H_2O diffusion in dacitic and andesitic melts. In: Xue X, Kanzaki M, Neuville DR, and Kawamoto T (eds.) *Structure and Properties of Silicate Melts and Fluids*, *Geochimica et Cosmochimica Acta* 68, pp. 5139–5150.

Bergantz GW and Breidenthal RE (2001) Non-stationary entrainment and tunneling eruptions: A dynamic link between eruption processes and magma mixing. *Geophysical Research Letters* 28(16): 3075–3078.

Bergantz GW and Ni J (1999) Numerical study of sedimentation by dripping instabilities in viscous fluids. *International Journal of Multiphase Flow* 25: 307–320.

Blake S and Fink JH (2000) On the deformation and freezing of enclaves during magma mixing. *Journal of Volcanology and Geothermal Research* 95: 1–8.

Blong RJ (1984) *Volcanic Hazards: A Sourcebook on the Effects of Eruptions*. FL, Orlando: Academic Press, 427 pp.

Bluth GJS, Rose WI, Sprod IE, and Krueger AJ (1997) Stratospheric loading of sulfur from explosive volcanic eruptions. *Journal of Geology* 105: 671–683.

Bonali FL, Tibaldi A, Corazzato C, Tormey DR, and Lara LE (2012) Quantifying the effect of large earthquakes in promoting eruptions due to stress changes on magma pathway: The Chile case. *Tectonophysics* 583: 54–67. http://dx.doi.org/10.1016/j.tecto.2012.10.025.

Brodsky EE (2004) International Association of volcanology and chemistry of the earth's interior. In: IUGG Working Group Geosciences (ed.) *The Future*, pp. 23–27. International Union of Geodesy and Geophysics. http://www.iugg.org/publications/reports/geosciences.pdf (accessed January 2013).

Brodsky EE and Prejean SG (2005) New constraints on mechanisms of remotely triggered seismicity at Long Valley Caldera. *Journal of Geophysical Research* 110: B04302. http://dx.doi.org/10.1029/2004JB003211.

Burgisser A, Arbaret L, Druitt TH, and Giachetti T (2011) Pre-explosive conduit conditions of the 1997 Vulcanian explosions at Soufrière Hills Volcano, Montserrat: II. Overpressure and depth distributions. *Journal of Volcanology and Geothermal Research* 199: 193–205.

Burgisser A and Gardner JE (2005) Experimental constraints on degassing and permeability in volcanic conduit flow. *Bulletin of Volcanology* 67: 42–56.

Bursik M (2001) Effect of wind on the rise height of volcanic plumes. *Geophysical Research Letters* 28: 3621–3624. http://dx.doi.org/10.1029/2001GL013393.

Cabato J, Costa F, and Newhall C (2012) The plumbing system of Mayon Volcano (Philippines): Insights from crystal zoning patterns and volatile contents (abstr). In: *Goldschmidt Conference 2012, Montreal*.

Carbone D and Poland MP (2012) Gravity fluctuations induced by magma convection at Kilauea Volcano, Hawaii. *Geology* 40: 4. http://dx.doi.org/10.1130/G33060.1.

Carey SN and Sigurdsson H (1989) The intensity of plinian eruptions. *Bulletin of Volcanology* 51: 28–40.

Carn SA, Krueger AJ, Bluth GJS, et al. (2003) Volcanic eruption detection by the Total Ozone Mapping Spectrometer (TOMS) instruments: A 22-year record of sulphur dioxide and ash emissions. In: Oppenheimer C, Pyle DM, and Barclay J (eds.) *Volcanic Degassing*, pp. 177–202. London: Geological Society of London, Special Publication 213.

Carn SA, Krueger AJ, Krotkov NA, Yang K, and Evans K (2009) Tracking volcanic sulphur dioxide clouds for aviation hazard mitigation. *Natural Hazards* 51: 325–343.

Carroll MR and Holloway JR (eds.) (1994) Reviews in mineralogy. In: *Volatiles in Magmas*, vol. 30, p. 517. Washington, DC, Mineralogical Society of America.

Cas RAF and Wright JV (1987) *Volcanic Successions: Ancient and Modern*. London: Allen and Unwin.

Casadevall TJ (ed.) (1994) *Volcanic ash and aviation safety*, In: *Proceedings of the First International Symposium on Volcanic Ash and Aviation Safety, US Geological Survey Bulletin 2047: 450*.

Cashman KV (2004) Volatile controls on magma ascent and eruption. In: Sparks RSJ and Hawkesworth CJ (eds.) *The State of the Planet: Frontiers and Challenges in Geophysics. Geophysical Monograph*, 150 Washington, DC: American Geophysical Union. http://dx.doi.org/10.1029/150GM10.

Cashman KV (2006) Bowen Lecture: Bubbles, crystals, and melts – 'the three-phase world of volcanic eruptions' (abs). *Eos, Transactions American Geophysical Union* 87(52): V21E-02, Fall Meeting Supplement.

Cashman KV and Mangan MT (1994) Physical aspects of magmatic degassing. Part II: Constraints on vesiculation processes from textural studies of eruptive products. In: Carroll MR and Holloway JR (eds.) *Volatiles in Magma. Reviews in Mineralogy* 30: pp. 447–478. Washington, DC: Mineralogical Society of America.

Cashman KV and Sparks RSJ (2013) How volcanoes work: A 25-year perspective. *Geological Society of America Bulletin*. http://dx.doi.org/10.1130/B30720.1.

Cashman KV, Sturtevant B, Papale P, and Navon O (2000) Magmatic fragmentation. In: Sigurdsson H, Houghton BF, McNutt SR, Rymer H, and Stix J (eds.) *Encyclopedia of Volcanoes*, pp. 421–430. San Diego, CA: Academic Press.

Castro JN and Dingwell DB (2009) Rapid ascent of rhyolitic magma at Chaitén volcano, Chile. *Nature* 461: 780–783. http://dx.doi.org/10.1038/nature08458.

Cervelli P (2000) DISLOC3D, Displacements and their derivatives from a dislocation in an elastic half-space (software based on Okada 1992). Version 1.0, September 22, 2000, http://www.avo.alaska.edu/Software/Kilauea/disloc3d.html (accessed January 2013).

Chen JK, Taylor FW, Edwards RL, Cheng H, and Burr GS (1995) Recent emerged reef terraces of the Yenkahe resurgent block, Tanna, Vanuatu: Implications for volcanic, landslide, and tsunami hazards. *Journal of Geology* 103: 577–590.

Chiodini G, Granieri D, Avino R, Caliro S, Costa A, and Werner C (2005) Carbon dioxide diffuse degassing and estimation of heat release from volcanic and hydrothermal systems. *Journal of Geophysical Research* 110: B08204. http://dx.doi.org/10.1029/2004JB003542.

Chouet B (1992) A seismic model for the source of long-period events and harmonic tremor. In: Gasparini P, Scarpa R, and Aki K (eds.) *Volcanic Seismology*, pp. 133–156. Berlin: Springer.

Chouet B (1996a) Long-period seismicity: Its source and use in eruption forecasting. *Nature* 380: 309–316.

Chouet BA (1996b) New methods and future trends in seismological volcano monitoring. In: Scarpa R and Tilling RI (eds.) *Monitoring and Mitigation of Volcano Hazards*, pp. 23–97. Berlin: Springer.

Cole SE, Cronin SJ, Sherburn S, and Manville V (2009) Seismic signals of snow-slurry lahars in motion: 25 September 2007, Mt Ruapehu, New Zealand. *Geophysical Research Letters* 36: L09405. http://dx.doi.org/10.1029/2009GL038030.

Coombs ML, Eichelberger JC, and Rutherford MJ (2000) Magma storage and mixing conditions for the 1953–1974 eruptions of Southwest Trident volcano, Katmai National Park, Alaska. *Contributions to Mineralogy and Petrology* 140: 99–118.

Coombs ML and Gardner JE (2004) Reaction rim growth on olivine in silicic melts; implications for magma mixing. *American Mineralogist* 89: 748–759.

Cooper KM and Reid MR (2003) Re-examination of crystal ages in recent Mount St. Helens Lavas: Implications for magma reservoir processes. *Earth and Planetary Science Letters* 213: 149–167.

Costa F, Dohmen R, and Chakraborty S (2008) Time scales of magmatic processes from modelling the zoning patterns of crystals. *Reviews in Mineralogy and Geochemistry* 69: 545–594.

Daag AS, Tubianosa BS, Newhall CG, et al. (1996) Monitoring sulfur dioxide emission at Mount Pinatubo. In: Newhall CG and Punongbayan RS (eds.) *Fire and Mud: Eruptions and Lahars of Mount Pinatubo, Philippines*, pp. 409–414. Quezon City and Seattle: Philippine Institute of Volcanology and Seismology and University of Washington Press.

Day SJ (1996) Hydrothermal pore fluid pressure and the stability of porous, permeable volcanoes. In: McGuire WJ, Jones AP, and Neuberg J (eds.) *Volcano Instability on the Earth and other Planets,* 110: pp. 77–93. London: Geological Society of London, Geological Society of London Special Publication.

Devine JD, Rutherford MJ, and Gardner JE (1997) Petrologic determination of ascent rates for the 1995–1997 Soufriere Hills Volcano andesitic magma. *Geophysical Research Letters* 25: 3673–3676.

Devine JD, Rutherford MJ, Norton GE, and Young SR (2003) Magma storage region processes inferred from geochemistry of Fe–Ti oxides in andesitic magma, Soufriere Hills Volcano, Montserrat, W.I. *Journal of Petrology* 44: 1375–1400.

Dieterich JH and Decker RW (1975) Finite element modeling of surface deformation associated with volcanism. *Journal of Geophysical Research* 80: 4094–4102.

Dingwell DB (1998) Recent experimental progress in the physical description of silicic magma relevant to explosive volcanism. In: Gilbert JS and Sparks RSJ (eds.) *The Physics of Explosive Eruption*, pp. 9–26. London: Geological Society of London Geological Society of London, Special Publication 145.

Dingwell DB (2003) Liquid to glass: Quantifying properties and structures of melts across the glass transition. In: De Vivo B and Bodnar RJ (eds.) *IAVCEI Developments in Volcanology 5: Melt Inclusions in Volcanic Systems: Methods, Applications, and Problems*, pp. 45–63. Amsterdam: Elsevier.

Dobran F (2001) *Volcanic Processes: Mechanisms in Material Transport*, 590 pp. New York: Academic/Plenum/Kluwer.

Doukas MP and Gerlach TM (1995) Sulfur dioxide scrubbing during the 1992 eruptions of Crater Peak, Mount Spurr Volcano, Alaska. In: Keith TEC (ed.) *The 1992 Eruptions of Crater Peak Vent, Mount Spurr Volcano, Alaska,* 2139: pp. 47–57. Denver, CO: US Geological Survey. US Geological Survey Bulletin.

Doyle EE, Cronin SJ, and Thouret J-C (2011) Defining conditions for bulking and debulking in lahars. *Geological Society of America Bulletin* 123: 1234–1246. http://dx.doi.org/10.1130/B30227.1.

Druitt TH, Costa F, Deloule E, Dungan M, and Scaillet B (2012) Decadal to monthly timescales of magma transfer and reservoir growth at a caldera volcano. *Nature* 482: 77–80. plus supplementary material.

Dunbar NW, Cashman KV, and Dupré R (1994) Crystallization processes of anorthoclase phenocrysts in the Mount Erebus magmatic system: Evidence from crystal composition, crystal size distributions, and volatile contents of melt inclusions. In: Kyle PR (ed.) *Volcanological and Environmental Studies of Mount Erebus, Antarctica. America Geophysical Union Antarctic Research Series,* pp. 129–146. Washington, DC: American Geophysical Union.

Dzurisin D (1980) Influence of fortnightly earth tides at Kilauea Volcano, Hawaii. *Geophysical Research Letters* 7: 925–928.

Dzurisin D (2003) A comprehensive approach to monitoring volcano deformation as a window on the eruption cycle (corrected). *Reviews of Geophysics* 41(2): 1009.

Dzurisin D (ed.) (2007) *Volcano Deformation: Geodetic Monitoring Techniques*, p. 441. Berlin: Springer-Praxis.

Dzurisin D, Lisowski M, and Wicks CW (2009) Continuing inflation at the Three Sisters volcanic center, central Oregon Cascade Range, USA, from GPS, leveling, and InSAR observations. *Bulletin of Volcanology* 71: 1091–1110.

Dzurisin D, Lisowski M, Wicks CW, Poland MP, and Endo ET (2006) Geodetic observations and modeling of magmatic inflation at the Three Sisters Volcanic Center, central Oregon Cascade Range, USA. *Journal of Volcanology and Geothermal Research* 150: 35–54.

Dzurisin D and Lu Z (2007) Interferometric synthetic-aperture radar (InSAR). In: Dzurisin D (ed.) *Volcano Deformation: Geodetic Monitoring Techniques*, pp. 153–194. Berlin: Springer-Praxis.

Eichelberger JC (1980) Vesiculation of mafic magma during replenishment of silicic magma reservoirs. *Nature* 288(5790): 446–450.

Eichelberger JC (1995) Silicic volcanism; ascent of viscous magmas from crustal reservoirs. *Annual Reviews of Earth and Planetary Sciences* 23: 41–63.

Eichelberger JC, Carrigan CR, Westrich HR, and Price RH (1986) Non-explosive silicic volcanism. *Nature* 323(6089): 598–602.

Emter D (1997) Tidal triggering of earthquakes and volcanic events. In: Wilhelm H, Zürn W, and Wenzel H-G (eds.) *Tidal Phenomena*, pp. 293–309. Berlin: Springer.

Ewert JW and Harpel CJ (2004) In harm's way: Population and volcanic risk. *Geotimes* 49: 14–17.

Fedotov SA, Gorelchik VI, and Stepanov VV (1980) Seismological studies on the mechanism of the Large Tolbachik fissure eruption, 1975–1976. *Bulletin of Volcanology* 43(1): 73–84.

Fehler M (1983) Observations of volcanic tremor at Mount St. Helens volcano. *Journal of Geophysical Research* 88: 3476–3484.

Feuillard M, Allegre CJ, Brandeis G, et al. (1983) The 1975–1977 crisis of La Soufrière de Guadeloupe (F.W.I.): A still-born magmatic eruption. *Journal of Volcanology and Geothermal Research* 16: 317–334.

Fischer TB, Arehart GB, Sturchio NC, and Williams SN (1996) The relationship between fumarole gas composition and eruptive activity at Galeras Volcano, Colombia. *Geology* 24: 531–534.

Fiske RS (1984) Volcanologists, journalists, and the concerned local public A tale of two crises in the eastern Caribbean. In: National Research Council Geophysics Study Committee (ed.) *Explosive Volcanism*, pp. 110–121. Washington, DC: National Academy Press.

Fitterman DV (1978) Electrokinetic and magnetic anomalies associated with dilatent regions in a layered earth. *Journal of Geophysical Research* 83: 5923–5928.

Fournier RO (1999) Hydrothermal processes related to movement of fluid from plastic into brittle rock in the magmatic-epithermal environment. *Economic Geology* 94: 1193–1211.

Fournier RO (2007) Hydrothermal systems and volcano geochemistry. In: Dzurisin D (ed.) *Volcano Deformation: Geodetic Monitoring Techniques*, pp. 323–341. Berlin: Springer.

Fournier d'Albe EM (1979) Objectives of volcano monitoring and prediction. *Journal of the Geological Society of London* 136: 321–326.

Francis P and Oppenheimer C (2004) *Volcanoes*, 2nd edn., 521 pp. Oxford, New York: Oxford Press.

Galle B, Johansson M, Rivera C, et al. (2010) Network for Observation of Volcanic and Atmospheric Change (NOVAC)—A global network for volcanic gas monitoring: Network layout and instrument description. *Journal of Geophysical Research, [Atmospheres]* 115: D5. http://dx.doi.org/10.1029/2009JD011823.

Gardner CA and Guffanti MC (2006) US Geological Survey's alert notification system for volcanic activity. US Geological Survey Fact Sheet 2006–3139, 4 pp, http://pubs.usgs.gov/fs/2006/3139/, accessed January 2013.

Gerlach TM, Doukas MP, McGee KA, and Kessler R (1998) Three-year decline of magmatic CO_2 emissions from soils of a Mammoth Mountain tree kill, Horseshoe Lake, California, 1995–1997. *Geophysical Research Letters* 25: 1947–1950.

Gerlach TM, Westrich HR, and Symonds RB (1996) Preeruption vapor in magma of the climactic Mount Pinatubo eruption: Source of the giant stratospheric sulfur dioxide cloud. In: Newhall CG and Punongbayan RS (eds.) *Fire and Mud: Eruptions and Lahars of Mount Pinatubo, Philippines*, pp. 415–433. Quezon City and Seattle: Philippine Institute of Volcanology and Seismology and University of Washington Press.

Giggenbach WF (1996) Chemical composition of volcanic gases. In: Scarpa R and Tilling RI (eds.) *Monitoring and Mitigation of Volcano Hazards*, pp. 221–256. Berlin: Springer.

Gil Cruz F and Chouet BA (1997) Long-period events, the most characteristic seismicity accompanying the emplacement and extrusion of a lava dome in Galeras Volcano, Colombia, in 1991. *Journal of Volcanology and Geothermal Research* 77: 121–158.

Glasby GP and Kasahara J (2001) Influence of tidal effects on the periodicity of earthquake activity in diverse geological settings with particular emphasis on submarine hydrothermal systems. *Earth-Science Reviews* 52(4): 261–297.

Gómez DM and Torres RA (1997) Unusual low-frequency volcanic seismic events with slowly decaying coda waves observed at Galeras and other volcanoes. *Journal of Volcanology and Geothermal Research* 77: 173–193.

Gonnermann HM and Manga M (2003) Explosive volcanism may not be an inevitable consequence of magma fragmentation. *Nature* 426: 432–435.

Goto A (1999) A new model for volcanic earthquake at Unzen Volcano: Melt Rupture Model. *Geophysical Research Letters* 26: 2541–2544. http://dx.doi.org/10.1029/1999GL900569.

Gottsmann J and Rymer H (2002) Deflation during caldera unrest: Constraints on subvolcanic processes and hazard prediction from gravity-height data. *Bulletin of Volcanology* 64: 338–348.

Gottsmann J, Rymer H, and Wooller LK (2005) On the interpretation of gravity variations in the presence of active hydrothermal systems: Insights from the Nisyros Caldera, Greece. *Geophysical Research Letters* 32: L23310. http://dx.doi.org/10.1029/2005GL024061.

Hammer JE, Cashman KV, and Voight B (2000) Magmatic processes revealed by textural and compositional trends in Merapi dome lavas. *Journal of Volcanology and Geothermal Research* 100: 165–192.

Harlow DH, Power JA, Laguerta EP, Ambubuyog G, White RA, and Hoblitt RP (1996) Precursory seismicity and forecasting of the June 15, 1991 eruption of Mount Pinatubo, Philippines. In: Newhall CG and Punongbayan RS (eds.) *Fire and Mud: Eruptions and Lahars of Mount Pinatubo, Philippines*, pp. 285–305. Quezon City and Seattle: Philippine Institute of Volcanology and Seismology and University of Washington Press.

Harrington RM and Brodsky EE (2007) Volcanic hybrid earthquakes that are brittle-failure events. *Geophysical Research Letters* 34: 6. http://dx.doi.org/10.1029/2006GL028714.

Hausback BP (2000) Geologic map of the Sasquatch Steps area, north flank of Mount St. Helens, Washington. US Geological Survey, Geologic Investigations Series, Report: I–2463, 1 sheet, http://geopubs.wr.usgs.gov/i-map/i2463 (accessed January 2013).

Hautmann S, Gottsmann J, Camacho AG, Fournier N, Sacks IS, and Sparks RSJ (2010) Mass variations in response to magmatic stress changes at Soufrière Hills Volcano, Montserrat (W.I.): Insights from 4-D gravity data. *Earth and Planetary Science Letters* 290: 83–89. http://dx.doi.org/10.1016/j.epsl.2009.12.004.

Hayashi Y and Morita Y (2003) An image of a magma intrusion process inferred from precise hypocentral migrations of the earthquake swarm east of the Izu Peninsula. *Geophysical Journal International* 153: 159–174.

Hayes SK, Montgomery DR, and Newhall CG (2002) Fluvial sediment transport and deposition following the 1991 eruption of Mount Pinatubo, Philippines. *Geomorphology* 45: 211–224.

Hellweg M (2003) The polarization of volcanic seismic signals: Medium or source? *Journal of Volcanology and Geothermal Research* 128: 159–176.

Hernandez PA, Notsu K, Salazar JM, et al. (2001) Carbon dioxide degassing by advective flow from Usu Volcano, Japan. *Science* 292: 83–86.

Hess K-U and Dingwell DB (1996) Viscosities of hydrous leucogranitic melts: A non-Arrhenian model. *American Mineralogist* 81: 1297–1300.

Hill DP (2006) Unrest in Long Valley Caldera, California 1978–2004. In: Troise C, De Natale G, and Kilburn CRJ (eds.) *Mechanisms of Activity and Unrest at Large Calderas*, vol. 269, pp. 1–24. London: Geological Society of London, Special Publication.

Hill DP, Langbein JO, and Prejean S (2003) Relations between seismicity and deformation during unrest in Long Valley Caldera, California, from 1995 through 1999. *Journal of Volcanology and Geothermal Research* 127: 175–193.

Hill DP, Pollitz F, and Newhall C (2002) Earthquake–volcano interactions. *Physics Today* 55(11): 41–47.

Hill DP and Prejean S (2007) *Dynamic Triggering, in Treatise on Geophysics*. 4. New York: Elsevier, 257–292 (revised, this volume).

Hincks TH, Aspinall W, and Woo G (2006) Evidence-based real-time hazard evaluation methodology and software for EXPLORIS volcanoes. Deliverable D6.4, EXPLORIS Project, 103 pp. (accessed from http://exploris.pi.ingv.it/confidential/products/Report/DELIVERABLES/PDF2006/EXPLORIS_D6.4.pdf, January 2013).

Hoblitt RP, Mori J, and Power JA (1996a) Computer visualization of earthquake hypocenters. In: Newhall CG and Punongbayan RS (eds.) *Fire and Mud: Eruptions and Lahars of Mount Pinatubo, Philippines*, pp. 383–385. Quezon City and Seattle: Philippine Institute of Volcanology and Seismology and University of Washington Press, + disk with Volquake program.

Hoblitt RP, Wolfe EW, Scott WE, Couchman MR, Pallister JS, and Javier D (1996b) The preclimactic eruptions of Mount Pinatubo, June 1991. In: Newhall CG and Punongbayan RS (eds.) *Fire and Mud: Eruptions and Lahars of Mount Pinatubo, Philippines*, pp. 457–511. Quezon City and Seattle: Philippine Institute of Volcanology and Seismology and University of Washington Press.

Ichihara M (2008) Dynamics of a spherical viscoelastic shell: Implications to a criterion for fragmentation/expansion of bubbly magma. *Earth and Planetary Science Letters* 265: 18–32.

Ichihara M, Rittel D, and Sturtevant B (2002) Fragmentation of a porous viscoelastic material: Implications to magma fragmentation. *Journal of Geophysical Research* 107(B10): ECV8-1–ECV8-14. http://dx.doi.org/10.1029/2001JB000591.

Iguchi M (1989) Distribution of the initial motions of volcanic microearthquakes (B-type) at Sakurajima Volcano. *Bulletin of the Disaster Prevention Research Institute, Kyoto University* 32(B-1): 13–22.

Iguchi M, Yakiwara H, Tameguri T, Hendrasto M, and Hirabayashi J (2008) Mechanism of explosive eruption revealed by geophysical observations at the Sakurajima, Suwanosejima and Semeru volcanoes. *Journal of Volcanology and Geothermal Research* 178: 1–9.

Ikuta R and Yamaoka K (2004) Temporal variation in the shear wave anisotropy detected using the Accurately Controlled Routinely Operated Signal System (ACROSS). *Journal of Geophysical Research* 109(B9): 14.

Ikuta R, Yamaoka K, Miyakawa K, Kunitomo T, and Kumazawa M (2002) Continuous monitoring of propagation velocity of seismic waves using ACROSS. *Geophysical Research Letters* 29(13): 4.

Ishibashi H (2009) Non-Newtonian behavior of plagioclase-bearing basaltic magma: Subliquidus viscosity measurement of the 1707 basalt of Fuji volcano, Japan. *Journal of Volcanology and Geothermal Research* 181: 78–88.

Ishihara K and Iguchi M (1989) The relationship between microearthquake swarms and volcanic activity at Sakurajima Volcano. *Bulletin of the Disaster Prevention Research Institute, Kyoto University* 32(B-1): 1–12.

Iverson RM (2008) Dynamics of seismogenic volcanic extrusion resisted by a solid surface plug, Mount St. Helens, 2004–2005. In: Sherrod DR, Scott WE, and Stauffer PH (eds.) *A Volcano Rekindled: The Renewed Eruption of Mount St. Helens 2004–2006*, pp. 4125–4460, US Geological Survey Professional Paper 1750.

Jackson DB (1988) Geoelectric observations (including the September 1982 summit eruption). In: Wolfe EW (ed.) *The Puu Oo Eruption of Kilauea Volcano, Hawaii: Episodes 1 Through 20, January 3, 1983, Through June 8, 1984*, pp. 237–251. Washington, DC: US Geological Survey, US Geological Survey Professional Paper 1463.

Jackson DB, Kauahikaua J, and Zablocki CJ (1985) Resistivity monitoring of an active volcano using the controlled-source electromagnetic technique, Kilauea, Hawaii. *Journal of Geophysical Research* 90: 12545–12555.

Jaupart C (1998) Gas loss from magmas through conduit walls. In: Gilbert JS and Sparks RSJ (eds.) *The Physics of Explosive Volcanic Eruptions*, 145, pp. 73–90. London: Geological Society of London, Geological Society, London, Special Publication.

Jaupart C and Allègre CJ (1991) Gas content, eruption rate and instabilities of eruption regime in silicic volcanoes. *Earth and Planetary Science Letters* 102: 413–429.

Jellinek AM and Bercovici D (2011) Seismic tremors and magma wagging during explosive volcanism. *Nature* 470: 522–525.

Jellinek AM, Manga M, and Saar MO (2004) Did melting glaciers cause volcanic eruptions in eastern California? Probing the mechanics of dike formation. *Journal of Geophysical Research* 109: B09206. http://dx.doi.org/10.1029/2004JB002978.

Johnson DJ (1987) Elastic and inelastic magma storage at Kilauea Volcano. In: Decker RW, Wright TL, and Stauffer PH (eds.) *Volcanism in Hawaii*, pp. 1297–1306. Washington, DC: US Government Printing Office, US Geological Survey Professional Paper.

Johnston MJS (1989) Review of magnetic and electric field effects near active faults and volcanoes in the USA. *Physics of the Earth and Planetary Interiors* 57: 47–63.

Julian BR (1994) Volcanic tremor: Nonlinear excitation by fluid flow. *Journal of Geophysical Research* 99: 11859–11877.

Julian BR (2000) Periodic doubling and other nonlinear phenomena in volcanic earthquakes and tremor. *Journal of Volcanology and Geothermal Research* 101: 19–26.

Kaizuka S, Kato S, Nagaoka S, and Miyauchi T (1985) Geomorphology of Iwo-jima and surrounding sea floor. *Journal of Geography (Tokyo)* 94: 424–436.

Kameda M, Kuribara H, and Ichihara M (2008) Dominant time scale for brittle fragmentation of vesicular magma by decompression. *Geophysical Research Letters* 35: 14. http://dx.doi.org/10.1029/2008GL034530.

Kamo K, Ishihara K, and Tahira M (1994) Infrasonic and seismic detection of explosive eruptions at Sakurajima Volcano, Japan, and the PEGASAS-VE early-warning system. In: Casadevall TJ (ed.) *US Geological Survey Bulletin 2047: Volcanic Ash and Aviation Safety*, pp. 357–365. Denver, CO: US Geological Survey.

Katsui Y and Katz HR (1967) Lateral fissure eruptions in the southern Andes of Chile. *Journal of Faculty of Science Hokkaido University* 13(4): 433–448.

Kawakatsu H and Yamamoto M (2015) Volcano seismology. In: Schubert G (ed.) *Treatise on Geophysics*, vol. 4, pp. 389–420. Oxford: Elsevier.

Kazahaya K, Shinohara H, and Saito G (1994) Excessive degassing of Izu-Oshima volcano: Magma convection in a conduit. *Bulletin of Volcanology* 56(3): 207–216.

Kazahaya K, Shinohara H, and Saito G (2002) Degassing process of Satsuma-Iwojima volcano, Japan: Supply of volatile components from a deep magma chamber. *Earth, Planets and Space* 54: 327–335.

Kazahaya K, Shinohara H, Uto K, et al. (2004) Gigantic SO_2 emission from Miyakejima Volcano, Japan, caused by caldera collapse. *Geology* 32: 425–428.

Kedar S, Kanamori H, and Sturtevant B (1998) Bubble collapse as the source of tremor at Old Faithful Geyser. *Journal of Geophysical Research* 103(B10): 24283–24299.

Kedar S, Sturtevant B, and Kanamori H (1996) The origin of harmonic tremor at Old Faithful Geyser. *Nature* 379: 708–711.

Keppler H (1999) Experimental evidence for the source of excess sulfur in explosive volcanic eruptions. *Science* 284: 1652–1654.

Kieffer S (1984) Seismicity of Old Faithful Geyser: An isolated source of geothermal noise and possible analogue of volcanic tremor. *Journal of Volcanology and Geothermal Research* 22: 59–96.

Klein FW, Koyanagi RY, Nakata JS, and Tanigawa WR (1987) The seismicity of Kilauea's magma system. In: Decker RW, Wright TL, and Stauffer PH (eds.) *Volcanism in Hawaii*, 1350, pp. 1019–1185. Washington, DC: US Government Printing Office, US Geological Survey Professional Paper.

Klug C and Cashman KV (1996) Permeability development in vesiculating magmas: Implications for fragmentation. *Bulletin of Volcanology* 58: 87–100.

Konstantinou KI and Schlindwein V (2002) Nature, wavefield properties and source mechanism of volcanic tremor: A review. *Journal of Volcanology and Geothermal Research* 119: 161–187.

Koyaguchi T and Blake S (1989) The dynamics of magma mixing in a rising magma batch. *Bulletin of Volcanology* 52: 127–137.

Koyama M and Yoshida Y (1994) Relationship between eruptive history and crustal stress field based on cumulative changes in magma discharge from volcanoes. *Bulletin of the Volcanological Society of Japan* 39: 177–190 (in Japanese with English abstract).

Koyanagi RY, Chouet B, and Aki K (1987) Origin of volcanic tremor in Hawaii. Part 1: Data from the Hawaiian Volcano Observatory 1969–1985. In: Decker RW, Wright TL, and Stauffer PH (eds.) *Volcanism in Hawaii*, 1350, pp. 1221–1258. Washington, DC: US Government Printing Office, US Geological Survey Professional Paper.

Kress V (1997) Magma mixing as a source for Pinatubo sulphur. *Nature* 389(6651): 591–593.

Kumagai H, Miyakawa K, Negishi H, Inoue H, Obara K, and Suetsugu D (2003) Magmatic dike resonances inferred from very-long-period seismic signals. *Science* 299: 2058–2061.

LaHusen RG (1996) Detecting Debris Flows Using Ground Vibrations: US Geol Surv Fact Sheet 236–96, http://vulcan.wr.usgs.gov/Projects/AFM/Publications/FS236-96/ (accessed January 2013).

LaHusen RG (2005) Acoustic Flow Monitor System – User Manual. US Geological Survey Open-File Report 02–429, 22 pp.

Lara LE, Naranjo JA, and Moreno H (2004) Rhyodacitic fissure eruption in Southern Andes (Cordón Caulle; 40.5S) after the 1960 (Mw 9.5) Chilean earthquake: A structural interpretation. *Journal of Volcanology and Geothermal Research* 138: 127–138.

Leet RC (1988) Saturated and subcooled hydrothermal boiling in groundwater flow channels as a source of harmonic tremor. *Journal of Geophysical Research* 93: 4835–4849.

Legrand D, Calahorrano A, Guillier B, et al. (2002) Stress tensor analysis of the 1998–1999 tectonic swarm of northern Quito related to the volcanic swarm of Guagua Pichincha volcano, Ecuador. *Tectonophysics* 344: 15–36.

Lenat J-F (1995) Geoelectrical methods in volcano monitoring. In: McGuire B, Kilburn CRJ, and Murray J (eds.) *Monitoring Active Volcanoes*, pp. 248–274. London: UCL Press.

Linde AT, Agustsson K, Sacks IS, and Stefansson R (1993) Mechanism of the 1991 eruption of Hekla from continuous borehole strain monitoring. *Nature* 365: 737–740.

Linde AT, Sacks IS, Johnston MJS, Hill DP, and Bilham RG (1994) Increased pressure from rising bubbles as a mechanism for remotely triggered seismicity. *Nature* 371: 408–410.

Lipman PW, Moore JG, and Swanson DA (1981) Bulging of the north flank before the May 18 eruption: Geodetic data. In: Lipman PW and Mullineaux DR (eds.) *The 1980 Eruptions of Mount St. Helens, Washington*, 1250, pp. 143–156. Denver, CO: US Geological Survey, US Geological Survey Professional Paper.

Lockwood JP and Hazlett RW (2010) *Volcanoes: Global Perspectives*. Wiley-Blackwell: Chichester, 539 pp.

Lu Z, Masterlark T, Dzurisin D, Rykhus R, and Wicks C Jr. (2003) Magma supply dynamics at Westdahl Volcano, Alaska, modeled from satellite radar interferometry. *Journal of Geophysical Research* 108: 2354. http://dx.doi.org/10.1029/2002JB002311.

Luhr JF, Carmichael IS, and Varekamp JC (1984) The 1982 eruption of El Chichón volcano, Chiapas, Mexico: Mineralogy and petrology of the anhydrite-bearing pumices. *Journal of Volcanology and Geothermal Research* 23: 69–108.

Maclennan J, Jull M, McKenzie D, Slater L, and Gronvold K (2002) The link between volcanism and deglaciation in Iceland. *Geochemistry, Geophysics, Geosystems* 3: 25.

Major JJ (2004) Posteruption suspended sediment transport at Mount St. Helens: Decadal–scale relationships with landscape adjustments and river discharges. *Journal of Geophysical Research* 109. http://dx.doi.org/10.1029/2002JF000010.

Malone SD (1983) Volcanic earthquakes: Examples from Mount St. Helens. In: Kanamori H and Boschi E (eds.) *Earthquakes: Observations, Theory and Interpretation*, pp. 436–455. Amsterdam: Elsevier.

Malone SD and Qamar A (1984) Repetitive microearthquakes as the source for volcanic tremor at Mount St. Helens. *Eos, Transactions American Geophysical Union* 65: 1001.

Manga M and Brodsky E (2006) Seismic triggering of eruptions in the far field: Volcanoes and geysers. *Annual Reviews of Earth and Planetary Sciences* 34: 263–291.

Manga M and Loewenberg M (2001) Viscosity of magmas containing highly deformable bubbles. *Journal of Volcanology and Geothermal Research* 105: 19–24.

Mangan M and Sisson TW (2000) Delayed, disequilibrium degassing in rhyolite magma; decompression experiments and implications for explosive volcanism. *Earth and Planetary Science Letters* 183: 441–455.

Marcial S, Melosantos AA, Hadley KC, LaHusen RG, and Marso JN (1996) Instrumental lahar monitoring at Mount Pinatubo. In: Newhall CG and Punongbayan RS (eds.) *Fire and Mud: Eruptions and Lahars of Mount Pinatubo, Philippines*, pp. 1015–1022. Quezon City and Seattle: Philippine Institute of Volcanology and Seismology and University of Washington Press.

Marsh BD (1981) On the crystallinity, probability of occurrence, and rheology of lava and magma. *Contributions to Mineralogy and Petrology* 78: 85–98.

Marsh BD (2000) Magma chambers. In: Sigurdsson H, Houghton BF, McNutt SR, Rymer H, and Stix J (eds.) *Encyclopedia of Volcanoes*, pp. 191–206. San Diego, CA: Academic Press.

Martel C, Dingwell DB, Spieler O, Pichavant M, and Wilke M (2000) Fragmentation of foamed silicic melts: An experimental study. *Earth and Planetary Science Letters* 178(1–2): 47–58.

Martin VM, Morgan DJ, Jerram DA, Caddick MJ, Prior DJ, and Davidson JP (2008) Bang! Month-scale eruption triggering at Santorini Volcano. *Science* 321: 1178. http://dx.doi.org/10.1126/science.1159584.

Marzocchi W and Bebbington M (2012) Probabilistic eruption forecasting at short and long time scales. *Bulletin of Volcanology*. http://dx.doi.org/10.1007/s00045-012-0633-x.

Marzocchi W, Neri A, Newhall CG, Papale P (2007) Probabilistic volcanic hazard and risk assessment. *EOS Trans. AGU* 88(32): 318 (full article pdf).

Marzocchi W, Newhall C, and Woo G (2012) The scientific management of volcanic crises. *Journal of Volcanology and Geothermal Research* 248: 181–189.

Marzocchi W, Sandri L, Gasparini P, Newhall C, and Boschi E (2004) Quantifying probabilities of volcanic events: The example of volcanic hazard at Mt. Vesuvius. *Journal of Geophysical Research* 109: B11201. http://dx.doi.org/10.1029/2004JB003155.

Marzocchi W, Sandri L, and Selva J (2010) BET_VH: A probabilistic tool for long-term volcanic hazard assessment. *Bulletin of Volcanology* 72: 705–716. http://dx.doi.org/10.1007/s00445-010-0357-8.

Marzocchi W and Woo G (2007) Probabilistic eruption forecasting and the call for an evacuation. *Geophysical Research Letters* 34: L22310.

Mastin LG (1994) Explosive tephra emissions at Mount St. Helens, 1989–1991: The violent escape of magmatic gas following storms? *Geological Society of America Bulletin* 106: 175–185.

Matsumoto N, Sato T, Matsushima N, Akita F, Shibata T, and Suzuki A (2002) Hydrological anomalies associated with crustal deformation before the 2000 eruption of Usu Volcano, Japan. *Geophysical Research Letters* 29(5): 1057. http://dx.doi.org/10.1029/2001GL013968.

McGee KA, Gerlach TM, Kessler R, and Doukas MP (2000) Geochemical evidence for a magmatic CO_2 degassing event at Mammoth Mountain, California, September–December 1997. *Journal of Geophysical Research* 105: 8447–8456.

McGuire B, Kilburn CRJ, and Murray J (eds.) (1995) *Monitoring Active Volcanoes*, 421 pp. London: UCL Press.

McKee CO, Wallace DA, Almond RA, and Talai B (1981) Fatal hydro-eruption of Karkar volcano in 1979: Development of a maar-like crater. In: Johnson R.W. (ed.) Geological Survey of Papua New Guinea Memoir 10: Cooke-Ravian Volume of Volcanological Papers, pp. 63–84. Port Moresby: Geological Survey PNG.

McNutt SR (1992) Volcanic tremor. In: Nierenberg WA (ed.) *Encyclopedia of Earth System Science*, vol. 4, pp. 417–425. San Diego, CA: Academic Press.

McNutt SR (1996) Seismic monitoring and eruption forecasting of volcanoes. In: Scarpa R and Tilling RI (eds.) *Monitoring and Mitigation of Volcano Hazards*, pp. 99–146. Berlin: Springer.

McNutt SR (2000a) Volcanic seismicity. In: Sigurdsson H, Houghton BF, McNutt SR, Rymer H, and Stix J (eds.) *Encyclopedia of Volcanoes*, pp. 1015–1033. San Diego, CA: Academic Press.

McNutt SR (2000b) Seismic monitoring. In: Sigurdsson H, Houghton BF, McNutt SR, Rymer H, and Stix J (eds.) *Encyclopedia of Volcanoes*, pp. 1095–1119. San Diego, CA: Academic Press.

McNutt SR (2005) Volcanic seismology. *Annual Reviews of Earth and Planetary Sciences* 33: 461–491.

McNutt SR and Beavan RJ (1981) Volcanic earthquakes at Pavlof Volcano correlated with the solid Earth tide. *Nature* 294(5842): 615–618.

McNutt SR and Beavan RJ (1987) Eruptions of Pavlof Volcano and their possible modulation by ocean load and tectonic stresses. *Journal of Geophysical Research* 92: 11509–11523.

Métrich N, Allard P, Aiuppa A, et al. (2011) Magma and volatile supply to post-collapse volcanism and block resurgence in Siwi Caldera (Tanna Island, Vanuatu arc). *Journal of Petrology* 52: 1077–1105.

Miller V and Savage M (2001) Changes in seismic anisotropy after volcanic eruptions: Evidence from Mount Ruapehu. *Science* 293: 2231–2233. http://dx.doi.org/10.1126/science.1063463.

Minakami T (1960) Fundamental research for predicting volcanic eruptions. Part I. Earthquakes and crustal deformations originating from volcanic activities. *Bulletin of Earthquake Research Institute* 38: 497–544.

Minakami T (1974) Seismology of volcanoes in Japan. In: Civetta L, Gasparini P, Luongo G, and Rapolla A (eds.) *Physical Volcanology*, pp. 1–27. Amsterdam: Elsevier.

Mizutani H, Ishido T, Yokokura T, and Ohnishi S (1976) Electrokinetic phenomena associated with earthquakes. *Geophysical Research Letters* 3: 365–368.

Mogi K (1958) Relations between the eruptions of various volcanoes and the deformations of the ground surfaces around them. *Bulletin of Earthquake Research Institute* 36: 99–134.

Molina I, Burgisser A, and Oppenheimer C (2012) Numerical simulations of convection in crystal-bearing magmas: A case study of the magmatic system at Erebus, Antarctica. *Journal of Geophysical Research* 117: B07209. http://dx.doi.org/10.1029/2011JB008760.

Moore PL, Iverson NR, and Iverson RM (2008) Frictional properties of the Mount. St. Helens gouge. In: Sherrod DR, Scott WE, and Stauffer PH (eds.) *A Volcano Rekindled: The Renewed Eruption of Mount St. Helens 2004–2006,* 1750, pp. 415–424. US Geological Survey Professional Paper.

Moore JG, Normark WR, and Holcomb RT (1994) Giant Hawaiian landslides. *Annual Reviews of Earth and Planetary Sciences* 22: 119–144.

Moran SC, Newhall C, and Roman DC (2011) Failed magmatic eruptions: Late-stage cessation of magma ascent. *Bulletin of Volcanology* 73: 115–122. http://dx.doi.org/10.1007/s0445-010-0444-x.

Mourtada-Bonnefoi CC and Laporte D (2004) Kinetics of bubble nucleation in a rhyolitic melt: An experimental study of the effect of ascent rate. *Earth and Planetary Science Letters* 218(3–4): 521–537. http://dx.doi.org/10.1016/S0012-821X(03)00684-8.

Mueller S, Scheu B, Spieler O, and Dingwell DB (2008) Permeability control on magma fragmentation. *Geology* 36(5): 399–402. http://dx.doi.org/10.1130/G24605A.1.

Nagaoka Y, Nishida K, Aoki Y, Takeo M, and Ohminato T (2012) Seismic imaging of magma chamber beneath an active volcano. *Earth and Planetary Science Letters* 333–334: 1–8.

Nakada S, Shimizu H, and Ohta K (1999) Overview of the 1990–1995 eruption at Unzen Volcano. *Journal of Volcanology and Geothermal Research* 89: 1–22.

Nakamichi H, Ukawa M, and Sakai S (2004) Precise hypocenter locations of midcrustal low-frequency earthquakes beneath Mt. Fuji, Japan. *Earth, Planets and Space* 56: e37–e40.

Nakamura K (1974) Preliminary estimate of global volcanic production rate. In: Colp JL and Furumoto AS (eds.) *The Utilization of Volcano Energy*, pp. 273–285. Albuquerque, NM, USA: Sandia Labs.

Nakao S, Morita Y, Goto K, et al. (2012) Crustal deformation after the eruption on January, 2011 by continuous GPS observation in Kirishima Volcano (abstr.). In: *Japan Geosciences Union Meeting, Chiba, Japan, May 2012* (manuscript in review, Earth Planets Space).

Neuberg J (2000) Characteristics and causes of shallow seismicity in andesitic volcanoes. *Philosophical Transactions of the Royal Society of London A* 358: 1533–1546.

Newhall CG (2000) Volcano warnings. In: Sigurdsson H, Houghton B, McNutt SR, Rymer H, and Stix J (eds.) *Encyclopedia of Volcanoes*, pp. 1185–1197. San Diego: Academic Press.

Newhall CG (2004) Degassing of arc volcanoes: Generalizations from WOVOdat (abstract) IAVCEI General Assembly, Pucon Chile, November 2004.

Newhall CG, Albano SE, Matsumoto N, and Sandoval T (2001) Roles of groundwater in volcanic unrest. *Journal of the Geological Society of the Philippines* 56: 69–84.

Newhall CG, Fujita E, Hill D, Kobayashi T, Lowenstern J, and Ukawa M (2003) Comparative seismicity, deformation, and magma supply at Iwo-jima, Long Valley, and Yellowstone calderas (abstract). In: *Abstracts, IUGG, Sapporo, July 2003*.

Newhall CG and Hoblitt RP (2002) Constructing event trees for volcanic crises. *Bulletin of Volcanology* 64: 3–20.

Newhall CG, Power JA, and Punongbayan RS (2002) Pinatubo: 'to make grow'. *Science* 295: 1241–1242.

Newhall CG, Punongbayan RS, and Gerlach TM (1994) Tight and leaky volcanoes: Implications for forecasting explosive eruptions. *Atti dei Convegni Lincei* 112: 13–21.

Newhall CG and Self S (1982) The Volcanic Explosivity Index (VEI): An estimate of explosive magnitude for historical volcanism. *Journal of Geophysical Research* 87: 1231–1238.

Nishimura T (2004) Pressure recovery in magma due to bubble growth. *Geophysical Research Letters* 31. http://dx.doi.org/10.1029/2004GL019810.

Nishimura T (2006) Ground deformation due to magma ascent with and without degassing. *Geophysical Research Letters* 33: L23309. http://dx.doi.org/10.1029/2006GL028101.

NOAA Satellite and Information Service (2013) Volcanic Ash Advisory Centers. URL: http://www.ssd.noaa.gov/VAAC/vaac.html (accessed January 2013).

Noguchi K and Kamiya H (1963) Prediction of volcanic eruption by measuring the chemical composition and amounts of gases. *Bulletin of Volcanology* 26: 367–378.

Okada Y (1992) Internal Deformation due to shear and tensile faults in a half-space. *Bulletin of the Seismological Society of America* 82: 1018–1049.

Okumura S, Nakamura M, Nakano T, Uesugi K, and Tsuchiyama A (2010) Shear deformation experiments on vesicular rhyolite: Implications for brittle fracturing, degassing, and compaction of magmas in volcanic conduits. *Journal of Geophysical Research* 115: B6. http://dx.doi.org/10.1029/2009JB006904.

Okumura S, Nakamura M, Takeuchi S, Tsuchiyama A, Nakano T, and Uesugi K (2009) Magma deformation may induce non-explosive volcanism via degassing through bubble networks. *Earth and Planetary Science Letters* 281: 267–274.

Okumura S, Nakamura N, Uesugi K, Nakano T, and Fujioka T (2013) Coupled effect of magma degassing and rheology on silicic volcanism. *Earth and Planetary Science Letters* 362: 163–170.

Oyagi N and Inokuchi T (1985) Geology of Iwo–jima. *Journal of Geography (Tokyo)* 94: 436–445.

Pal R (2003) Rheological behavior of bubble-bearing magmas. *Earth and Planetary Science Letters* 207: 165–179.

Pallister JS, Hoblitt RP, Meeker GP, Knight RJ, and Siems DF (1996) Magma mixing at Mount Pinatubo: Petrographic and chemical evidence from the 1991 deposits. In: Newhall CG and Punongbayan RS (eds.) *Fire and Mud: Eruptions and Lahars of Mount Pinatubo, Philippines*, pp. 687–731. Quezon City and Seattle: Philippine Institute of Volcanology and Seismology and University of Washington Press.

Pallister JS, Major JJ, Pierson TC, et al. (2010) Interdisciplinary studies of Eruption at Chaitén Volcano, Chile. *Eos, Transactions American Geophysical Union* 91(42): 381–382. http://dx.doi.org/10.1029/2010EO420001.

Pallister JS, Thornber CR, Cashman KV, et al. (2008) Petrology of the 2004–2006 Mount St. Helens lava dome: Implications for magma plumbing and eruption triggering. In: Sherrod DR, Scott WE, and Stauffer PH (eds.) *A Volcano Rekindled: The Renewed Eruption of Mount St. Helens 2004–2006,* 1750, pp. 647–702, US Geological Survey Professional Paper.

Pérez NM, Hernández PA, Padilla G, et al. (2011) Global CO_2 emission from volcanic lakes. *Geology* 39: 235–239.

Phillips JC, Lane SJ, Lejeune A-M, and Hilton M (1995) Gum rosin–acetone system as an analogue to the degassing behaviour of hydrated magmas. *Bulletin of Volcanology* 57(4): 263–268.

Pierson TC and Costa JE (1987) A rheologic classification of subaerial sediment-water flows. *Reviews in Engineering Geology* 7: 1–12.

Polacci M, Papale P, and Rosi M (2001) Textural heterogeneities in pumices from the climactic eruption of Mount Pinatubo, 15 June 1991, and implications for magma ascent dynamics. *Bulletin of Volcanology* 63: 83–97.

Power JA, Stihler SD, White RA, and Moran SC (2004) Observations of deep long-period (DLP) seismic events beneath Aleutian arc volcanoes; 1989–2002. *Journal of Volcanology and Geothermal Research* 138: 243–266.

Power JA, White RA, Murray TL, et al. (2001) Seismicity and forecasting of the 1991 eruption of Mount Pinatubo: A ten-year retrospective (abs). *Eos, Transactions American Geophysical Union,* Fall Meeting, Abstract U31A–03.

Prejean SG and Brodsky EE (2011) Volcanic plume height measured by seismic waves based on a mechanical model. *Journal of Geophysical Research* 116: B01306. http://dx.doi.org/10.1029/2010JB007620.

Pyle DM (2000) Sizes of volcanic eruptions. In: Sigurdsson H, Houghton BF, McNutt SR, Rymer H, and Stix J (eds.) *Encyclopedia of Volcanoes*, pp. 263–269. San Diego, CA: Academic Press.

Ramos EG, Hamburger MW, Pavlis GL, and Laguerta EP (1999) The low-frequency earthquake swarms at Mount Pinatubo, Philippines; implications for magma dynamics. *Journal of Volcanology and Geothermal Research* 92: 295–320.

Reid ME (2004) Massive collapse of volcano edifices triggered by hydrothermal pressurization. *Geology* 32: 373–376.

Roeloffs EA and Linde AT (2007) Borehole observations of continuous strain and fluid pressure. In: Dzurisin D (ed.) *Volcano Deformation: Geodetic Monitoring Techniques*, pp. 305–322. Berlin: Springer, chapter 9.

Roman DC, Neuberg J, and Luckett RR (2006) Assessing the likelihood of volcanic eruption through analysis of volcanotectonic earthquake fault-plane solutions. *Earth and Planetary Science Letters* 248: 244–252.

Romero R (2004) The International Airways Volcano Watch (IAVW). In: *Proceedings Second International Conference on Volcanic Ash and Aviation Safety, Alexandria, VA June 2004, Session 4*, p. 10.

Rowe CA, Thurber CH, and Power JA (2001) Precise correlation-based phase repacking of Pinatubo pre-eruption seismicity illuminates families of repeating earthquakes and a refined image of precursory hypocenter locations (abs). *Eos, Transactions American Geophysical Union* 82(47): U32A-0002, Fall Meeting Supplement, Abstract.

Rust AC, Manga M, and Cashman KV (2003) Determining flow type, shear rate and shear stress in magmas from bubble shapes and orientations. *Journal of Volcanology and Geothermal Research* 122: 111–132.

Rutherford MJ and Devine JD (1996) Preeruption pressure and temperature conditions and volatiles in the 1991 dacitic magma of Mount Pinatubo. In: Newhall CG and Punongbayan RS (eds.) *Fire and Mud: Eruptions and Lahars of Mount Pinatubo, Philippines*, pp. 751–766. Quezon City and Seattle: Philippine Institute of Volcanology and Seismology and University of Washington Press.

Rutherford MJ and Gardner JE (2000) Rates of magma ascent. In: Sigurdsson H, Houghton BF, McNutt SR, Rymer H, and Stix J (eds.) *Encyclopedia of Volcanoes*, pp. 207–217. San Diego, CA: Academic Press.

Ryan MP (1987) Neutral buoyancy and the mechanical evolution of magmatic systems. In: Mysen BO (ed.) *Magmatic Processes: Physicochemical Principles*, vol. 1, pp. 259–287. St Louis, MO: The Geochemical Society, Special Publication, Geochemical Society.

Rymer H (1996) Microgravity monitoring. In: Scarpa R and Tilling RI (eds.) *Monitoring and Mitigation of Volcano Hazards*, pp. 169–197. Berlin: Springer.

Rymer H and Williams-Jones G (2000) Volcanic eruption prediction; magma chamber physics from gravity and deformation measurements. *Geophysical Research Letters* 27(16): 2389–2392.

Sanchez JJ, McNutt SR, Power JA, and Wyss M (2004) Spatial variations in the frequency-magnitude distribution of earthquakes at Mount Pinatubo volcano. *Bulletin of the Seismological Society of America* 94: 430–438.

Sasai Y, Shimomura T, Hamano Y, et al. (1990) Volcano-magnetic effect observed during the 1986 eruption of Izu-Oshima volcano. *Journal of Geomagnetism and Geoelectricity* 42: 291–317.

Sato T, Matsumoto N, Takahashi M, et al. (2001) Observation of water level, temperature and electric conductivity of groundwater around Usu Volcano. In: Tomiya A (ed.) *The 2000 eruption of Usu Volcano*, vol. 52, pp. 143–148. Tokyo: Geological Survey of Japan, Bulletin of the Geological Society of Japan.

Scarpa R and Tilling RI (eds.) (1996) *Monitoring and Mitigation of Volcano Hazards*, 841 pp. Berlin: Springer.

Schmincke H-U (2004) *Volcanism*, 324 pp. Berlin: Springer.

Segall P (2010) *Earthquake and Volcano Deformation*, 432 pp. Princeton: Princeton University Press.

Shaw HR (1972) Viscosities of magmatic silicate liquids; an empirical method of prediction. *American Journal of Science* 272(9): 870–893.

Shaw HR (1985) Links between magma-tectonic rate balances, plutonism, and volcanism. *Journal of Geophysical Research* 90(B13): 11275–11288.

Sherrod DR, Scott WE, and Stauffer PH, eds (2008) A Volcano Rekindled: The Renewed Eruption of Mount St. Helens 2004–2006, US Geological Survey Professional Paper 1750.

Shibata T and Akita F (2001) Precursory changes in well water prior to the March 2000 eruption of Usu volcano, Japan. *Geophysical Research Letters* 28(9): 1799–1802.

Shimomura Y, Nishimura T, and Sato H (2006) Bubble growth processes in magma surrounded by an elastic medium. *Journal of Volcanology and Geothermal Research* 155: 307–322.

Shimozuru D (1972) A seismological approach to the prediction of volcanic eruptions. In: UNESCO (ed.) *The Surveillance and Prediction of Volcanic Activity*, pp. 19–45. Paris: UNESCO.

Shinohara H (2008) Excess degassing from volcanoes and its role on eruptive and intrusive activity. *Reviews of Geophysics* 46: RG4005. http://dx.doi.org/10.1029/2007RG000244.

Siebert L (1996) Hazards of large volcanic debris avalanches and associated eruptive phenomena. In: Scarpa R and Tilling RI (eds.) *Monitoring and Mitigation of Volcano Hazards*, pp. 541–572. Berlin: Springer.

Siebert L, Glicken H, and Ui T (1987) Volcanic hazards from Bezymianny- and Bandai-type eruptions. *Bulletin of Volcanology* 49: 435–459.

Siebert L, Simkin T, and Kimberly P (2010) *Volcanoes of the World*, 3rd edn., 551 pp. Washington, DC/Berkeley, CA: Smithsonian Institution/University of California Press.

Sigurdsson H, Houghton BF, McNutt SR, Rymer H, and Stix J (eds.) (2000) *Encyclopedia of Volcanoes*, 1417 pp. San Diego, CA: Academic Press.

Simkin T (1993) Terrestrial volcanism in space and time. *Annual Reviews of Earth and Planetary Sciences* 21: 427–452.

Simkin T and Siebert L (2000) Earth's volcanoes and eruptions: An overview. In: Sigurdsson H, Houghton BF, McNutt SR, Rymer H, and Stix J (eds.) *Encyclopedia of Volcanoes*, pp. 249–261. San Diego, CA: Academic Press.

Sparks RSJ (1978) The dynamics of bubble formation and growth in magmas: A review and analysis. *Journal of Volcanology and Geothermal Research* 3: 1–37.

Sparks RSJ (2003) Dynamics of magma degassing. In: Oppenheimer C, Pyle DM, and Barclay J (eds.) *Volcanic Degassing*, vol. 213, pp. 5–22. London, Geological Society of London, Geological Society of London, Special Publication.

Sparks RSJ and Marshall LA (1986) Thermal and mechanical constraints on mixing between mafic and silicic magmas. *Journal of Volcanology and Geothermal Research* 29: 99–124.

Sparks RSJ, Sigurdsson H, and Wilson L (1977) Magma mixing: A mechanism for triggering acid explosive eruptions. *Nature* 267: 315–318.

Sparks RSJ and Young SR (2002) The eruption of Soufrière Hills Volcano, Montserrat (1995–1999); overview of scientific results. In: Druitt TH and Kokelaar BP (eds.) *Geological Society of London, Memoir*, vol. 21, pp. 45–69. Bath, UK, Geological Society Publishing House.

Spera FJ (1984) Carbon dioxide in petrogenesis. Part III: Role of volatiles in the ascent of alkaline magma with special reference to xenolith-bearing magmas. *Contributions to Mineralogy and Petrology* 88: 217–232.

Spera FJ (2000) Physical properties of magmas. In: Sigurdsson H, Houghton BF, McNutt SR, Rymer H, and Stix J (eds.) *Encyclopedia of Volcanoes*, pp. 171–190. San Diego, CA: Academic Press.

Stasiuk VM, Barclay J, Carroll MR, et al. (1996) Degassing during magma ascent in the Mule Creek vent (USA). *Bulletin of Volcanology* 58: 117–130.

Stevenson DS and Blake S (1998) Modelling the dynamics and thermodynamics of volcanic degassing. *Bulletin of Volcanology* 60: 307–317.

Stoiber RE and Jepsen A (1973) Sulfur dioxide contributions to the atmosphere by volcanoes. *Science* 182: 577–579.

Stoiber RE, Malinconico LL, and Williams SN (1983) Use of the correlation spectrometer at volcanoes. In: Tazieff H and Sabroux J-C (eds.) *Forecasting Volcanic Events*, pp. 425–444. New York: Elsevier.

Swanson DA, Casadevall TJ, Dzurisin D, Malone SD, Newhall CG, and Weaver CS (1983) Predicting eruptions at Mount St Helens June 1980 through December 1982. *Science* 221: 1369–1376.

Symonds RB, Gerlach TM, and Reed MH (2001) Magmatic gas scrubbing: Implications for volcano monitoring. *Journal of Volcanology and Geothermal Research* 108: 303–341.

Symonds RB, Rose WI, Bluth GJS, and Gerlach TM (1994) Volcanic-gas studies: Methods, results, and applications. In: Carroll MR and Holloway JR (eds.) *Reviews in Mineralogy. Volatiles in Magmas*, vol. 30, pp. 1–66. Washington, DC: Mineralogical Society of America.

Takada A (1999) Variations in magma supply and magma partitioning: The role of tectonic settings. *Journal of Volcanology and Geothermal Research* 93(1–2): 93–110.

Thomas ME and Neuberg J (2012) What makes a volcano tick—A first explanation of deep multiple seismic sources in ascending magma. *Geology* 40: 351–354. http://dx.doi.org/10.1130/G32868.1.

Tilling RI (1989) *Volcanic Hazards, Short Course in Geology*, vol. 1, 123 pp. Washington, DC: American Geophysical Union.

Tilling RI (1991) Reducing volcanic risk: Are we winning some battles but losing the war? *Earthquakes and Volcanoes* 22(3): 133–137.

Toramaru A (2006) BND (bubble number density) decompression rate meter for explosive volcanic eruptions. *Journal of Volcanology and Geothermal Research* 154: 303–316.

Toramaru A, Noguchi S, Oyoshihara S, and Tsune A (2008) MND (microlite number density) water exsolution rate meter. *Journal of Volcanology and Geothermal Research* 175: 156–167.

Trunk L and Bernard A (2008) Investigating crater lake warming using ASTER thermal imagery: Case studies at Ruapehu, Poás, Kawah Ijen, and Copahué. *Journal of Volcanology and Geothermal Research* 178: 259–270.

Tsuruga K, Yomogida K, Honda S, Ito H, Ohminato T, and Kawakatsu H (1997) Spatial and temporal variations of volcanic earthquakes at Sakurajima Volcano, Japan. *Journal of Volcanology and Geothermal Research* 75: 337–358.

Tsuya H (1955) Geological and petrological studies of volcano Fuji. Part 5: On the 1707 eruption of volcano Fuji. *Bulletin of Earthquake Research Institute, University of Tokyo* 33: 341–383.

Tuffen H and Dingwell DB (2004) Fault textures in volcanic conduits: Evidence for seismic trigger mechanisms during silicic eruptions. *Bulletin of Volcanology* 67: 370–387. http://dx.doi.org/10.1007/s00445-004-0383-5.

Tungol NM (2002) Engineering Geology of the Pasig-Potrero Lahar Fan, Eastern Mount Pinatubo, Luzon Island, Philippines. PhD Thesis, University of Canterbury, Christchurch.

Tungol NM and Regalado MTM (1996) Rainfall, acoustic flow monitor records, and observed lahars of the Sacobia River in 1992. In: Newhall CG and Punongbayan RS (eds.) *Fire and Mud: Eruptions and Lahars of Mount Pinatubo, Philippines*, pp. 1023–1032. Quezon City and Seattle: Philippine Institute of Volcanology and Seismology and University of Washington Press.

Ui T (1983) Volcanic dry avalanche deposits—Identification and comparison with nonvolcanic debris stream deposits. *Journal of Volcanology and Geothermal Research* 18: 135–150.

Ukawa M (2005) Deep low-frequency earthquake swarm in the mid crust beneath Mount Fuji (Japan) in 2000 and 2001. *Bulletin of Volcanology* 68: 17. http://dx.doi.org/10.1007/s00445-005-0419-5.

Ukawa M, Fujita E, Ueda H, Kumagai T, Nakajima H, and Morita H (2005) Long-term geodetic measurements of large scale deformation at low-jima caldera, Japan. *Journal of Volcanology and Geothermal Research* 150: 21. http://dx.doi.org/10.1016/j.jvolgeores.2005.07.008.

UNESCO (1972) *The Surveillance and Prediction of Volcanic Activity*. Paris: UNESCO.

Van Vliet-Lanoë B, Bourgeois O, Dauteuil O, Embry JC, Schneider JL, and Guillou H (2005) Deglaciation and volcano-seismic activity in northern Iceland: Holocene and early Eemian. *Geodinamica Acta* 18: 81–100.

Varekamp JC and Rowe GL (2000) Crater lakes. *Journal of Volcanology and Geothermal Research* 97(1–4): 508.

Vergniolle S and Jaupart C (1986) Separated two-phase flow and basaltic eruptions. *Journal of Geophysical Research* 91(12): 12842–12860.

Voight B and Elsworth D (1997) Failure of volcanic slopes. *Geotechnique* 47: 1–31.

Voight B, Sparks RSJ, Miller AD, et al. (1999) Magma flow instability and cyclic activity at Soufrière Hills Volcano, Montserrat, British West Indies. *Science* 283: 1138–1142.

Voight B, Young KD, Hidayat D, et al. (2000) Deformation and seismic precursors to dome-collapse and fountain-collapse nuées ardentes at Merapi Volcano, Java, Indonesia, 1994–1998. *Journal of Volcanology and Geothermal Research* 100: 261–287.

Waite GP, Nadeau PA, and Lyons JJ (2013) Variability in eruption style and associated very-long-period events at Fuego volcano, Guatemala. *Journal of Geophysical Research, Solid Earth* http://dx.doi.org/10.1002/jgrb.50075.

Walker GPL (1973) Explosive volcanic eruptions: A new classification scheme. *Geologische Rundschau* 62: 431–446.

Walker GPL (1980) The Taupo Pumice: Product of the most powerful known (ultraplinian) eruption? *Journal of Volcanology and Geothermal Research* 8: 69–94.

Wallace PJ (2001) Volcanic SO_2 emissions and the abundance and distribution of exsolved gas in magma bodies. *Journal of Volcanology and Geothermal Research* 108: 85–106.

Wallace P and Anderson AT Jr (2000) Volatiles in magmas. In: Sigurdsson H, Houghton BF, McNutt SR, Rymer H, and Stix J (eds.) *Encyclopedia of Volcanoes*, pp. 149–170. San Diego, CA: Academic Press.

Walter TR (2007) How a tectonic earthquake may wake up volcanoes: Stress transfer during the 1996 earthquake-eruption sequence at the Karymsky volcano group, Kamchatka. *Earth and Planetary Science Letters* 264: 347–359.

Walter TR and Amelung F (2007) Volcanic eruptions following $M \geq 9$ megathrust earthquakes: Implications for the Sumatra–Andaman volcanoes. *Geology* 35: 539–542.

Watt SFL, Pyle DM, and Mather TA (2013) The volcanic response to deglaciation: Evidence from glaciated arcs and a reassessment of global eruption records. *Earth-Science Reviews* 122: 77–102. http://dx.doi.org/10.1016/j.earscirev.2013.03.007.

White RA (1996) Precursory deep long-period earthquakes at Mount Pinatubo: Spatio-temporal link to a basalt trigger. In: Newhall CG and Punongbayan RS (eds.) *Fire and Mud: Eruptions and Lahars of Mount Pinatubo, Philippines*. Quezon City and Seattle: Philippine Institute of Volcanology and Seismology and University of Washington Press.

White SM, Crisp JA, and Spera FJ (2006) Long-term volumetric eruption rates and magma budgets. *Geochemistry, Geophysics, Geosystems* 7(3): 20. http://dx.doi.org/10.1029/2005GC001002.

White RA and Power JA (2001) Distal volcano-tectonic earthquakes (DVT's): Diagnosis and use in eruption forecasting (abs.). *Eos, Transactions American Geophysical Union* 82(47): U32A-0001, Fall Meeting Supplement.

Wicks CW Jr, Dzurisin D, Ingebritsen S, Thatcher W, Lu Z, and Iverson J (2002) Magmatic activity beneath the quiescent: Three Sisters volcanic center, central Oregon Cascade Range, USA. *Geophysical Research Letters* 29: 7. http://dx.doi.org/10.1029/2001GL014205.

Wilson L (1976) Explosive volcanic eruptions. Part III: Plinian eruption columns. *Geophysical Journal of the Royal Astronomical Society* 45: 543–556.

Wilson T, Cole J, Cronin S, Stewart C, and Johnston D (2011) Impacts on agriculture following the 1991 eruption of Vulcan Hudson, Patagonia: Lessons for recovery. *Natural Hazards* 57: 185–212.

Wilson L, Sparks RSJ, and Walker GPL (1980) Explosive volcanic eruptions. Part IV: The control of magma properties and conduit geometry on eruption column behavior. *Geophysical Journal of the Royal Astronomical Society* 63: 117–148.

Wilson TM, Stewart C, Sword-Daniels V, et al. (2012) Volcanic ash impacts on critical infrastructure. *Physics and Chemistry of the Earth* 45–46: 5–23.

Winson A, Costa F, Newhall C, and Woo G (2012) How successful are volcano alert levels at anticipating eruptions? (abstr 3P2.1-16), Cities on Volcanoes 7, Colima Mexico November 2012.

Witham F (2011) Conduit convection, magma mixing, and melt inclusion trends at persistently degassing volcanoes. *Earth and Planetary Science Letters* 301: 345–352.

Witter JB (2003) Convection of Magma in Volcanic Conduits as a Degassing Mechanism at Active Volcanoes. PhD Thesis, University of Washington.

Witter JB, Kress VC, Delmelle P, and Stix J (2004) Volatile degassing, petrology, and magma dynamics of the Villarrica Lava Lake, Southern Chile. *Journal of Volcanology and Geothermal Research* 134: 303–337.

Witter JB, Kress VC, and Newhall CG (2005) Volcán Popocatépetl, Mexico. Petrology, magma mixing, and immediate sources of volatiles for the 1994-present eruption. *Journal of Petrology* 46: 2337–2366.

Woo G (2008) Probabilistic criteria for volcano evacuation decision. *Natural Hazards* 45: 87–97.

Woodhouse MJ, Hogg AJ, Phillips JC, and Sparks RSJ (2013) Interaction between volcanic plumes and wind during the 2010 Eyjafjallajökull eruption, Iceland. *Journal of Geophysical Research* 118: 1–18. http://dx.doi.org/10.1029/2012JB009592.

Wright JV, Smith AL, and Self S (1980) A working terminology of pyroclastic deposits. *Journal of Volcanology and Geothermal Research* 8: 315–336.

Zapata JA, Calvache ML, Cortes GP, et al. (1997) SO_2 fluxes from Galeras Volcano, Colombia, 1989–1995; progressive degassing and conduit obstruction of a Decade Volcano. *Journal of Volcanology and Geothermal Research* 77: 195–208.

Zhang Y and Behrens H (2000) H_2O diffusion in rhyolitic melts and glasses. *Chemical Geology* 169: 243–262.

Zhang Y, Xu Z, and Liu Y (2003) Viscosity of hydrous rhyolitic melts inferred from kinetic experiments, and a new viscosity model. *American Mineralogist* 88: 1741–1752.

Zlotnicki J (1995) Geomagnetic surveying methods. In: McGuire B, Kilburn CRJ, and Murray J (eds.) *Monitoring Active Volcanoes*, pp. 275–300. London: UCL Press.

Zlotnicki J, Sasai Y, Yvetot P, et al. (2003) Resistivity and self-potential changes associated with volcanic activity: The July 8, 2000 Miyake–jima eruption (Japan). *Earth and Planetary Science Letters* 205: 139–154.

Zobin V (2011) *Introduction to Volcanic Seismology*, 2nd edn. Amsterdam: Elsevier.

This page intentionally left blank

4.15 Volcano Seismology

H Kawakatsu, The University of Tokyo, Tokyo, Japan
M Yamamoto, Tohoku University, Sendai, Miyagi, Japan

© 2015 Elsevier B.V. All rights reserved.

4.15.1	**Introduction**	389
4.15.2	**Volcanic Seismic Signals**	390
4.15.2.1	Signals Observed by Short-Period Seismometers	390
4.15.2.1.1	Deep (volcanic) LF signals	391
4.15.2.2	Broadband Signals	391
4.15.2.2.1	Global observation	392
4.15.3	**Description of Volcanic Seismic Sources**	392
4.15.3.1	Equivalent Forces of General Seismic Sources	392
4.15.3.2	The Seismic Moment Tensor	393
4.15.3.2.1	Spherical source	395
4.15.3.2.2	Cylinder	396
4.15.3.2.3	Tensile crack	396
4.15.3.2.4	Compensated linear vector dipole	396
4.15.3.2.5	Determination and interpretation of a moment tensor	396
4.15.3.3	The Single Force	397
4.15.3.4	Waveform Analysis of Volcanic Seismic Sources	398
4.15.4	**Physical Mechanisms for Volcanic Seismic Signals**	399
4.15.4.1	Slow Waves in Solid/Fluid Composite	399
4.15.4.2	Resonating Sources: A Crack	400
4.15.4.3	Other Resonating Sources	401
4.15.4.4	Flow-Induced Oscillation	402
4.15.4.4.1	In fluid/gas two-phase system	403
4.15.4.5	Bubble Dynamics	403
4.15.5	**Observation and Analysis Aspects**	404
4.15.5.1	Broadband Seismometry	404
4.15.5.2	Array Analysis	405
4.15.6	**Models for Volcanic Seismic Signals**	405
4.15.6.1	Aso	406
4.15.6.2	Kilauea, Hawaii	408
4.15.6.3	Miyake	409
4.15.6.4	Stromboli	410
4.15.7	**Other Volcano-Specific Issues**	412
4.15.7.1	Explosion Quakes	412
4.15.7.1.1	Scaling law of volcanic explosion	413
4.15.7.2	Triggered Seismicity	413
4.15.7.3	Seismic Wave Scattering in Volcanoes	414
4.15.8	**Concluding Remarks**	414
Acknowledgments		415
References		415

4.15.1 Introduction

Volcano seismology is a field of volcanology in which seismological techniques are employed to help understand the physical conditions and dynamic states of volcanic edifices and volcanic fluid systems to such a level that it eventually contributes to predictions of initiation and cessation of hazardous volcanic activities. Seismology has been a powerful tool for this purpose, especially in the past 20–25 years as new observational and analysis techniques in seismology have become available (e.g., Chouet, 1996a,b; McNutt, 2005); they include digital and broadband seismometry and array and moment tensor waveform analyses. While the subjects that should be covered under the title of volcano seismology are quite broad and diverse, this chapter, however, neither intends to nor is capable to cover all of its areas; instead, we try to focus on some of seismological phenomena that appear to be specific to volcanoes. In a sense, volcano seismology here is a field of seismology that deals with problems specific to volcanic areas. Some of the subjects, such as seismic structure of volcanoes (tomography) and eruption monitoring (seismicity), are therefore intentionally omitted. The readers who are interested in

these subjects are advised to refer to the excellent reviews by Chouet (1996a,b, 2003) and McNutt (1996, 2002, 2005).

4.15.2 Volcanic Seismic Signals

For seismologists who have only observed seismograms of ordinary faulting earthquakes, the most intriguing aspect of volcano seismology may be the ample variety of waveforms recorded in volcanic environments. Such volcanic seismic signals range from totally chaotic continuous vibrations lasting for minutes, hours, or sometimes days to purely monotonic vibrations decaying with constant damping rates; they often show several spectral peaks whose frequency may vary during the course of volcanic activities, possibly reflecting the change of physical conditions of volcanic fluid systems; when eruptions occur, a variety of long-period (LP) signals up to as long as a few hundred seconds are generated and sometimes observed at seismic stations around the world. In the following, some of such intriguing volcanic seismic signals will be presented.

Terminology. Based on their appearance on short-period seismometers, volcanic seismic signals have been traditionally classified into four types: high-frequency or A type, low-frequency (LF) or B type, explosion quakes, and volcanic tremors (e.g., Minakami, 1974; for the detail of different terminology, see table 1 of McNutt, 1996). In this chapter, following this tradition, we use high frequency for frequencies higher than 5 Hz and low frequency for frequencies between 5 and 0.5 Hz. On the other hand, the usage of period remains a complicated issue in volcano seismology; the most widely accepted notation to describe volcanic seismic signals is that of Chouet (1996b), in which events of about 1 s are called 'LP' events and those with longer periods are called 'very-long-period' (VLP) and 'ultra-long-period' events. This definition, however, contradicts with the conventional usage of period in traditional earthquake seismology, where periods longer than the dominant period of the ambient microseismic noise (around 5–10 s) are called 'long-period' (Aki and Lee, 2003; McNutt, 1996, 2005). In this chapter, as we try to cover the subject in a wider context of seismology, we follow the convention of earthquake seismology and use the term LP for periods longer than 5 s and short-period for periods shorter than or around 1 s; for events with a period longer than 50 s, we may also use VLP. To avoid confusion, whenever appropriate, we will try to associate actual number in seconds to describe a period, for example, LP (10 s).

4.15.2.1 Signals Observed by Short-Period Seismometers

The conventional seismometry at active volcanoes, except for a few rare cases, has been conducted using short-period seismometers. Even with such band restricted sensors, wide varieties of waveforms recorded in volcanoes have led scientists to wonder what may be the origins of those signals. **Figure 1** shows a classical yet the most fascinating one of those recordings of volcanic tremors observed at Sakurajima Volcano, Japan (Kamo et al., 1977). Volcanic tremors are those sustained ground vibrations commonly observed at volcanoes, typically lasting for longer than 1 min, up to hours or days; they often show periodic spectral features as seen here with the lowest peak frequency around 1 Hz. During the course of activities, relative spectral content and peak frequencies may vary from time to time.

Figure 2 shows examples from Redoubt Volcano, Alaska, where different types of events are displayed: from the top,

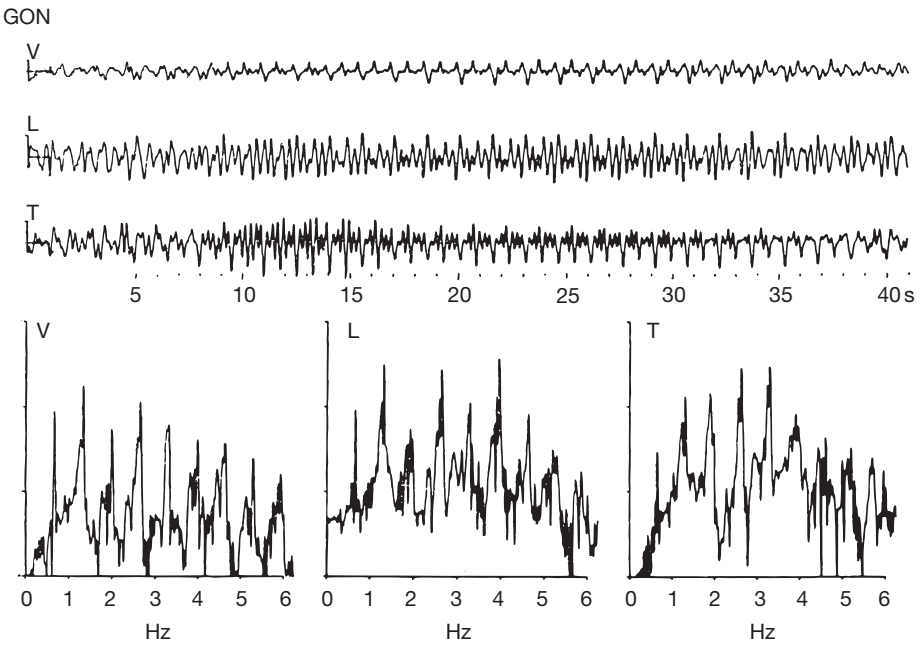


Figure 1 Volcanic tremor observed at Sakurajima. Top three traces show 40 s long seismograms for vertical, longitudinal, and transverse components. Corresponding spectra are shown below (or bottom). Adapted from Kamo K, Furuzawa T, and Akamatsu J (1977) Some natures of volcanic micro-tremors at the Sakurajima volcano. *Bulletin of the Volcanological Society of Japan* 22: 41–58.

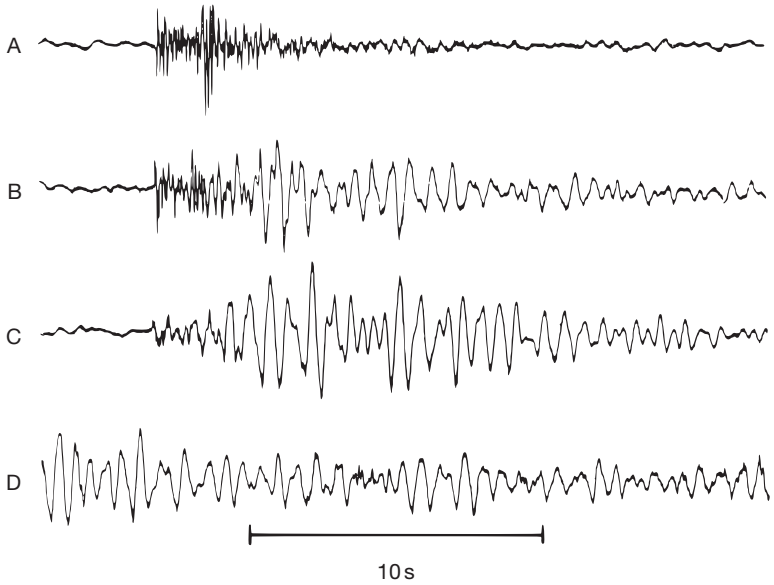


Figure 2 Typical waveforms observed at Redoubt. Adapted from McNutt SR (1996) Seismic monitoring and eruption forecasting of volcanoes: A review of the state-of-the-art and case histories. In: *Monitoring and Mitigation of Volcano Hazards*, pp. 99–146. New York: Springer-Verlag.

(a) high-frequency or volcano-tectonic earthquake, which most likely represents an ordinary brittle failure on a fault; (b) hybrid or mixed-frequency event; (c) LF event with frequency around 1 Hz; and (d) volcanic tremor. Note that the volcanic tremor here appears more random compared to those in Figure 1 and also its frequency content seems similar to that of the LF event (c). It is sometimes considered that tremors and LF events share a common excitation mechanism.

LF events in some volcanoes show characteristic similar spindle-shaped waveforms, suggesting a possible common origin for these events in different volcanoes.

Explosion quakes are observed when explosive eruptions occur and often associated with a high-frequency arrival on seismograms due to a shock wave originated at the vents. Figure 3 shows series of such events recorded in Langila Volcano, Papua New Guinea (Mori et al., 1989), sorted by the ratio of ground-transmitted explosion-wave and air-wave amplitudes. As those without an air-wave are classified as LF events, this figure also indicates that explosion quakes and LF events may share a common origin.

As noted earlier, the spectral content of these LF volcanic seismic signals may vary throughout the course of activities. Figure 4 shows such an example from Montserrat Volcano, West Indies; the temporal change of the system prior to an explosion is clearly manifested.

It appears clear that the LF nature of these seismograms is the most characteristic of volcanic seismic signals. Although they are characterized here just in terms of the frequency content, there exist similarities and diversities among them; their origins may be different from a volcano to another and may change from time to time even in a single volcano. Understanding of their origins and monitoring of their appearances to infer physical conditions and dynamic states of volcanic edifices and volcanic fluid systems are interesting and challenging tasks.

4.15.2.1.1 Deep (volcanic) LF signals

Most of volcanic seismic signals originate within a shallow part of the crust (<5 km). There are, however, certain types of volcanic seismic signals that originate either from the lower crust or from the uppermost mantle (15–50 km) right beneath active volcanoes. As they show similar LF content as volcanic tremors, they are often called deep LF volcanic tremors (or events). They are now reported in many volcanoes around the world (e.g., Aki and Koyanagi, 1981; Hasegawa and Yamamoto, 1994; Nakamichi et al., 2003; Pitt and Hill, 1994; White, 1996). As deep LF volcanic tremors may be related with the deep magma-supplying system of volcanoes, understanding of their origin seems extremely important for the long-term prediction of the activities of volcanoes.

4.15.2.2 Broadband Signals

Availability of broadband seismic records has brought wider perspectives in our view of volcanic activities. Figure 5 shows a very good example observed at Satsuma-Iwojima Volcano, Japan (Ohminato and Ereditato, 1997). The short-period sensor recording at ST3 shows amplitude modulation of high-frequency signal with a periodicity of 46–50 min. When this amplitude becomes small, the low-pass-filtered broadband record at ST1 shows LP (10 s) spike-like signals. This synchronicity of two different band signals continues, while the exact periodicity may fluctuate. Such an interplay of different band signals forces us to view the origin of these signals as a system (i.e., volcanic fluid system).

Figure 6 is another such example of a broadband record of a small phreatic eruption observed at Aso Volcano, Japan (Kaneshima et al., 1996; Kawakatsu et al., 2000). While the integrated displacement record exhibits a VLP (>100 s) signal corresponding to the inflation and deflation stages of a

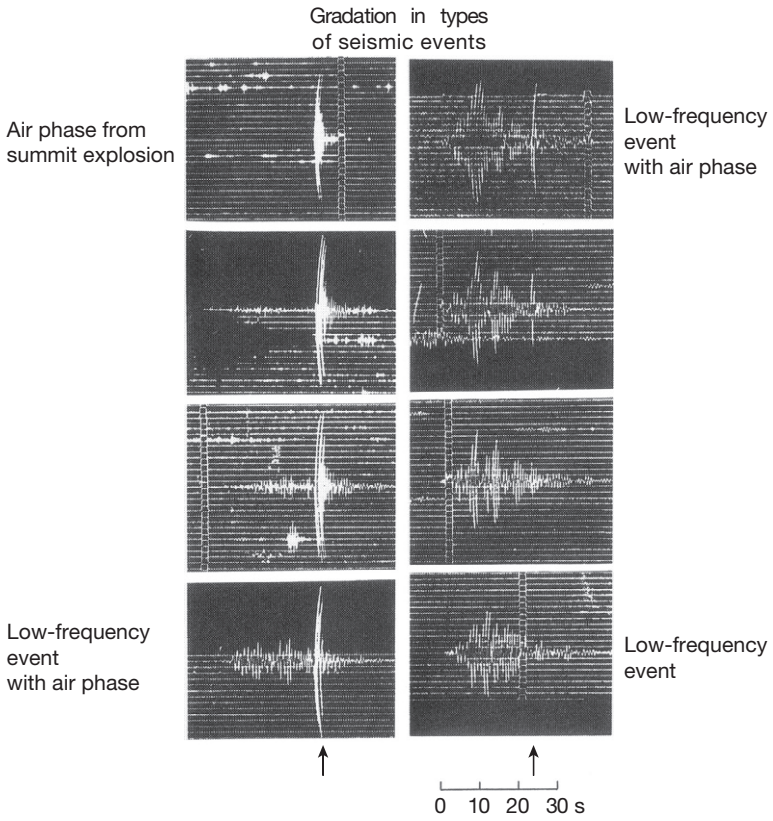


Figure 3 Range of volcanic seismic signals from the high-frequency air-wave phase to LF event observed at Langila. Arrows indicate timings of air-wave phases. Adapted from Mori J, Patia H, McKee C, et al. (1989) Seismicity associated with eruptive activity at Langila volcano, Papua New Guinea. *Journal of Volcanology and Geothermal Research* 38: 243–255.

crack-like conduit (Yamamoto et al., 1999) before and during the eruption, the raw velocity record exhibits short-period signals apparently due to the quick flow of volcanic fluids only seen in the deflation stage (i.e., eruption). Again, signals seen in two different bands allow us to vividly image the process of the eruption of a volcanic fluid system. Different physics apply at different frequencies, and this makes volcanic seismic signals essentially broadband.

4.15.2.2.1 Global observation

The last example (**Figure 7**) comes from the recordings of VLP (200–400 s) seismic signals generated by the eruption of Pinatubo Volcano, Philippines, observed at broadband stations around the world (Kanamori and Mori, 1992; Widmer and Zürn, 1992). The LP disturbances of the atmosphere generated by the eruption clouds are transmitted back to the solid Earth through the coupling between the two systems. Two spectral peaks observed in the seismic records are well explained by a theory that treats the solid Earth and the atmosphere as a single system (Watada and Kanamori, 2010); such a coupling between the solid and fluid parts of the Earth is also observed in the so-called background free oscillations of the Earth (Nishida et al., 2000). The broadband recordings of the Pinatubo eruption appear to have accelerated the emergence of a new class of seismology that treats the solid Earth, atmosphere, and ocean as a single system (cf. **Chapter 4.14**).

4.15.3 Description of Volcanic Seismic Sources

The diversity of observed volcanic seismic signals partly reflects the wide variety of seismic source geometries and dynamics present in active volcanoes. A general kinematic description of seismic sources in terms of 'moment tensor' and 'single force' has become common in volcano seismology, and the waveform inversion solving for them has become a general tool to understand the generation mechanisms of LP signals obtained with broadband instruments.

4.15.3.1 Equivalent Forces of General Seismic Sources

The concept of the seismic moment tensor was first introduced by Gilbert (1970) as a description of seismic sources more general than the conventional double-couple description. Later, Backus and Mulcahy (1976) gave a physical basis for such a general description within the framework of linear elastodynamics by introducing concepts of indigenous source and stress glut. An indigenous source is a seismic source, which originates within the Earth, and its equivalent force system exerts neither total force nor total torque at any instance to the Earth. Seismologists, however, soon realized that there exist such indigenous seismic events, which require single-force equivalent force systems (e.g., Hasegawa and Kanamori, 1987; Kanamori and Given, 1982; Kanamori et al., 1984;

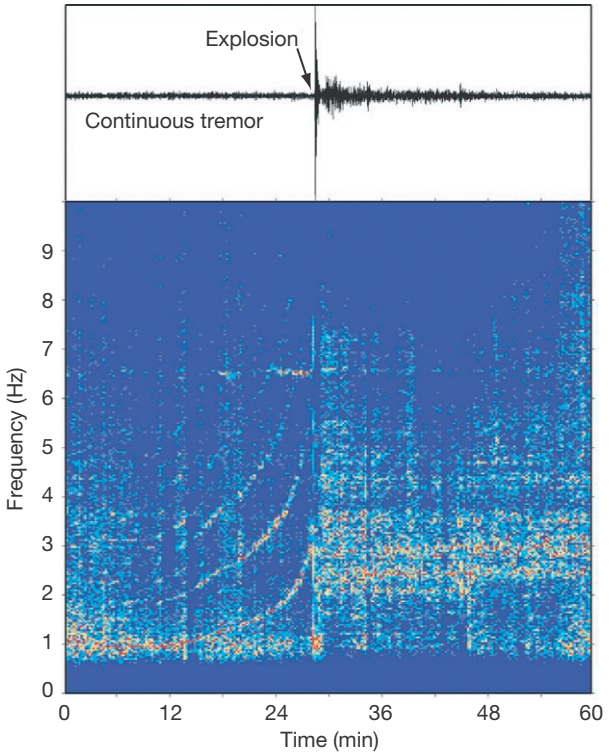


Figure 4 Example of tremor at Montserrat before an explosion. Adapted from Jousset P, Neuberg J, and Sturton S (2003) Modelling the time-dependent frequency content of low-frequency volcanic earthquakes. *Journal of Volcanology and Geothermal Research* 128: 201–223.

Kawakatsu, 1989); most of them happened to occur around active volcanoes. To circumvent this apparent conflict, Takei and Kumazawa (1994) extended Backus and Mulcahy's treatment to include fluid-dynamic description with a mass advection term and introduced a concept of inertial glut; they further gave a clear definition of single-force (and torque) source as a momentum (and angular momentum) exchange between the seismic source volume and the rest of the Earth. In the following, we briefly summarize their treatment.

To make the argument simple, we omit the Earth's self-gravity in the following; although gravity often takes the central role generating single forces in the actual Earth, this simplification does not alter the essence of the argument presented below (or in the following). The true Eulerian equation of motion of the Earth without external force may be written as

$$\rho^{\mathrm{t}} D_{\mathrm{t}}^2 u_j = \partial_i S_{ij} \qquad [1]$$

where $\rho^{\mathrm{t}}(x,t)$ denotes the true density, $u_j(x,t)$ is the displacement, $S_{ij}(x,t)$ is the true stress field, and D_{t} is the particle derivative $D_{\mathrm{t}} = \partial_{\mathrm{t}} + v_k \partial_k$ with $v_k(x,t)$ as the velocity. We compare this true equation of motion with the linear elastodynamic equation of motion

$$\rho^{\mathrm{m}} \partial_{\mathrm{t}}^2 u_j = \partial_i \varphi_{ij} \qquad [2]$$

where $\rho^{\mathrm{m}}(x,t)$ and $\varphi_{ij}(x,t)$, respectively, denote the model density and model stress field due to the linear stress–strain constitution equation (Hooke's law), which we use to describe the wave propagation in the Earth. The equivalent body force $\gamma_j^{\mathrm{V}}(x,t)$ in the linear elastodynamic system arises as the difference of the two equations [1] and [2] as

$$\rho^{\mathrm{m}} \partial_{\mathrm{t}}^2 u_j = \partial_i \varphi_{ij} + \gamma_j^{\mathrm{V}} \qquad [3]$$

where

$$\gamma_j^{\mathrm{V}} = \gamma_j^{\mathrm{S}} + \gamma_j^{\mathrm{D}} \qquad [4]$$

with

$$\gamma_j^{\mathrm{S}} = -\partial_i \Gamma_{ij} = -\partial_i \left(\varphi_{ij} - S_{ij} \right) \qquad [5]$$

$$\gamma_j^{\mathrm{D}} = \rho^{\mathrm{m}} \partial_{\mathrm{t}}^2 u_j - \rho^{\mathrm{t}} D_{\mathrm{t}}^2 u_j \qquad [6]$$

$\Gamma_{ij}(x,t) = \varphi_{ij}(x,t) - S_{ij}(x,t)$ in eqn [5] is the stress glut introduced by Backus and Mulcahy (1976), which represents failure of Hooke's law due to some nonlinear effect in the source region and is zero outside of the source region. The second term in the right hand of eqn [4] $\gamma_j^{\mathrm{D}}(x,t)$ is named as inertial glut by Takei and Kumazawa (1994) and is the difference of the inertial forces in the actual value and in the model, which is again zero outside of the source region. It can be shown that total force and total torque due to $\gamma_j^{\mathrm{S}}(x,t)$ integrated over the volume are zero at any instance, while those due to the inertial glut $\gamma_j^{\mathrm{D}}(x,t)$ are necessarily not.

The spatial integral of eqn [6] over the whole space

$$\gamma_j^{(0)\mathrm{D}} = \int \left(\rho^{\mathrm{m}} \partial_{\mathrm{t}}^2 u_j - \rho^{\mathrm{t}} D_{\mathrm{t}}^2 u_j \right) \mathrm{d}V \qquad [7]$$

can be shown to be

$$\gamma_j^{(0)\mathrm{D}} = -\partial_{\mathrm{t}}^2 \int_{V_{\mathrm{s}}} (\rho^{\mathrm{t}} - \rho^{\mathrm{m}}) u_j \mathrm{d}V \qquad [8]$$

which shows that $\gamma_j^{(0)\mathrm{D}}(t)$ originates from a difference between ρ^{t} and ρ^{m} in the source region V_{s} and thus may not vanish. On the other hand, it is shown that the time integral of $\gamma_j^{(0)\mathrm{D}}(t)$ during the event vanishes, that is,

$$\int_0^\infty \gamma_j^{(0)\mathrm{D}} \mathrm{d}t = 0 \qquad [9]$$

the requirement that any indigenous source has to satisfy. The authors further defined single-force and torque sources as linear and angular momenta exchange between the source region and the rest of the Earth and showed that, in the LP limit, the single-force source is equal to $\gamma_j^{(0)\mathrm{D}}(t)$.

4.15.3.2 The Seismic Moment Tensor

The spatial integral of stress glut is the seismic moment tensor M_{ij}, which describes the overall feature of the seismic source. When the wavelength of observed seismic waves is much longer than the spatial extent of the source, we may approximate the source by a point source moment tensor. Each component of M_{ij} corresponds to one set of opposing forces (dipole or force couple; Figure 8). The seismic moment tensor is a symmetrical second-order tensor and thus has six independent components, while a double-couple equivalent body force for a shear dislocation has only four degrees of freedom. These two extra degrees of freedom in the moment tensor are called non-double-couple

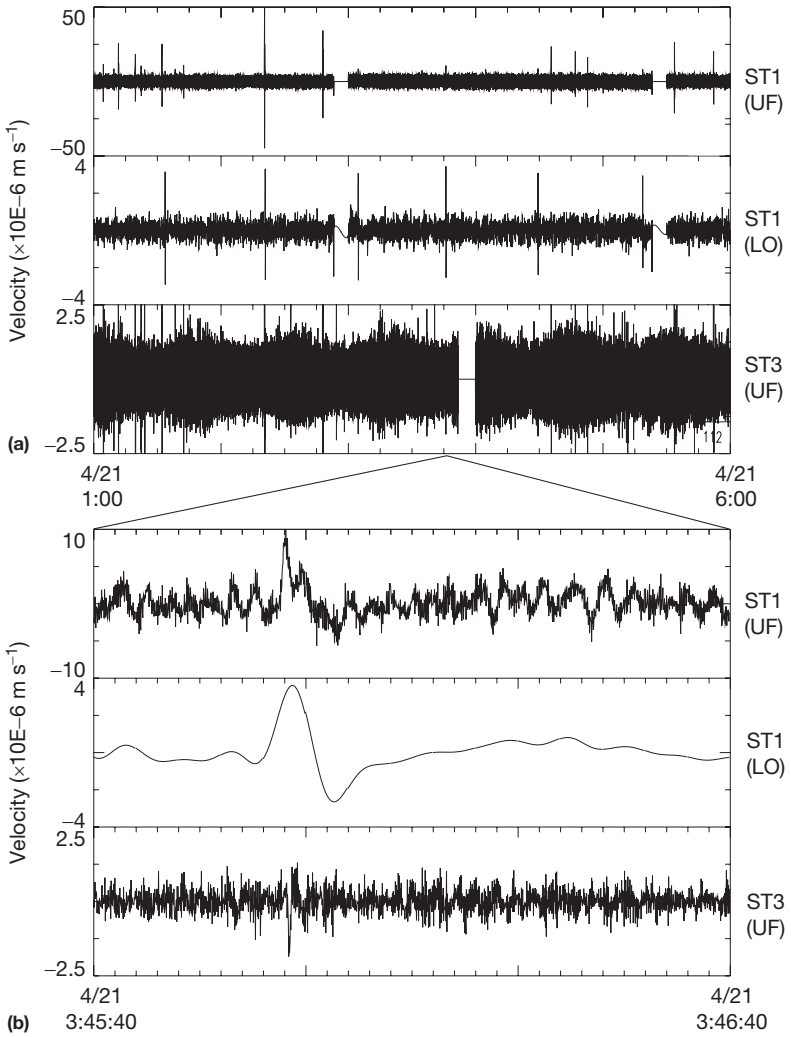


Figure 5 (a) Five-hour records of vertical components of velocity at Satsuma-Iwojima. A trace low-pass-filtered (LO) at 0.2 Hz is shown in between the two unfiltered records. (b) Expansion of the 1 min portion of the top panel. Adapted from Ohminato T and Ereditato D (1997) Broadband seismic observations at Satsuma-Iwojima volcano, Japan. *Geophysical Research Letters* 24: 2845–2848.

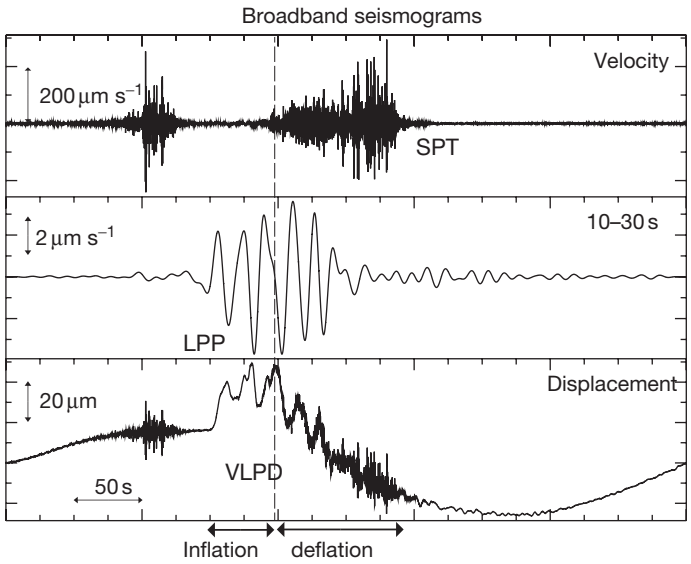


Figure 6 Broadband waveforms of a phreatic eruption observed at Aso. Adapted from Kawakatsu H, Kaneshima S, Matsubayashi H, et al. (2000) Aso94: Aso seismic observation with broadband instruments. *Journal of Volcanology and Geothermal Research* 101: 129–154.

components. For the general discussion on non-double-couple components, readers may refer to the following review articles (Julian et al., 1998; Miller et al., 1998). Seismic sources observed in the volcanic environments often show non-double-couple components. The end-members of such non-double-couple sources are discussed below (or in the following).

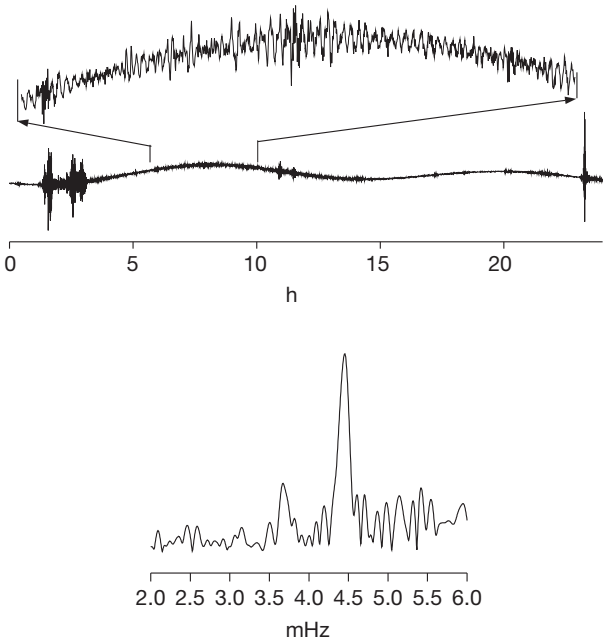


Figure 7 Broadband record of the Pinatubo eruption in 1991 observed in Japan. Adapted from Watada S and Kanamori H (2010) Acoustic resonant oscillations between the atmosphere and the solid earth during the 1991 Mt. Pinatubo eruption. *Journal of Geophysical Research* 115: B12319.

4.15.3.2.1 Spherical source

Let us consider an initial condition that a reservoir contains a certain amount of volcanic fluid (V) under a static pressure P. When a sudden volume increase of ΔV (measured under the same pressure condition) is introduced in the reservoir, either due to an injection of a new volcanic fluid or a thermal expansion, the reservoir should expand and act as a seismic source. This volumetric increase ΔV corresponds to the stress-free volumetric strain introduced by Eshelby (1957) (Aki and Richards, 2002, p. 53) and characterizes the amount of the increase of the volcanic fluid. On the other hand, due to the confining pressure of the surrounding elastic medium (matrix), the actual volumetric increase of the reservoir ΔV_m may be smaller than ΔV, the ratio $\Delta V_m/\Delta V$ depending on the geometry of the reservoir, with a resultant pressure increase ΔP_m.

The corresponding moment tensor for a spherical reservoir is given as

$$M = \Delta V \begin{pmatrix} \lambda + \frac{2}{3}\mu & 0 & 0 \\ 0 & \lambda + \frac{2}{3}\mu & 0 \\ 0 & 0 & \lambda + \frac{2}{3}\mu \end{pmatrix} \quad [10]$$

where the ratio

$$\frac{\Delta V_m}{\Delta V} = \frac{\lambda + \frac{2}{3}\mu}{\lambda + 2\mu} \quad [11]$$

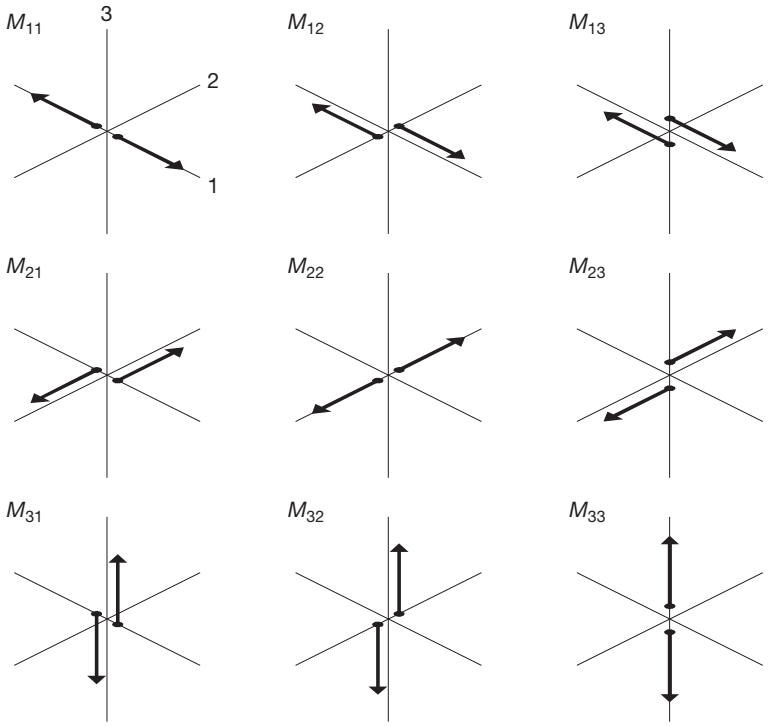


Figure 8 Nine force couples corresponding to moment tensor components.

and the actual volumetric strain and pressure increase may be related as

$$\Delta P_{\mathrm{m}} = \frac{4}{3}\mu\frac{\Delta V_{\mathrm{m}}}{V} \qquad [12]$$

(Aki and Richards, 2002, p. 61). As the actual volumetric change ΔV_{m} is also the volume change often used in geodetic analysis (e.g., Mogi, 1958), we propose to call this 'Mogi volume' to distinguish it from the stress-free volume increase ΔV.

4.15.3.2.2 Cylinder

A moment tensor corresponding to the radial expansion of a cylinder whose symmetry axis is x_1 may be given as

$$M = \Delta V\begin{pmatrix}\lambda & 0 & 0\\ 0 & \lambda+\mu & 0\\ 0 & 0 & \lambda+\mu\end{pmatrix} \qquad [13]$$

where

$$\frac{\Delta V_{\mathrm{m}}}{\Delta V} = \frac{\lambda+\mu}{\lambda+2\mu} \qquad [14]$$

and

$$\Delta P_{\mathrm{m}} = \mu\frac{\Delta V_{\mathrm{m}}}{V} \qquad [15]$$

4.15.3.2.3 Tensile crack

A moment tensor corresponding to an opening of a thin crack in the direction of x_1-axis may be given as

$$M = \Delta V\begin{pmatrix}\lambda+2\mu & 0 & 0\\ 0 & \lambda & 0\\ 0 & 0 & \lambda\end{pmatrix} \qquad [16]$$

In case of a thin crack, two volumetric changes turn out to be the same. The Mogi volume here is given as $\Delta V = \Delta V_{\mathrm{m}} = S\cdot\Delta u$, where S and Δu are the crack area and opening width, respectively, and it is often referred to as 'potency' in seismological literatures (e.g., Ben-Menahem and Singh, 1981).

4.15.3.2.4 Compensated linear vector dipole

The isotropic components of the moment tensors of these three volumetric sources, when expressed in terms of stress-free volumetric change ΔV, are the same:

$$I = \frac{1}{3}\mathrm{trace}\,(M) = \left(\lambda+\frac{2}{3}\mu\right)\Delta V = \kappa\Delta V \qquad [17]$$

where κ is the bulk modulus of the elastic matrix. When a movement of magma from a reservoir to another excites seismic waves, such a seismic source should be observed as a summation of two volumetric sources with opposite signs. Equation [17] indicates that the corresponding moment tensor has zero isotropic component called compensated linear vector dipole by Knopoff and Randall (1970) (**Figure 9**).

4.15.3.2.5 Determination and interpretation of a moment tensor

Due to structural complexities of volcanic edifices, the conventional first motion polarity analysis method to determine source mechanisms of volcanic events is unlikely to give reliable estimates of moment tensors. Although the amplitude of radiated seismic waves contains more information, often, the absolute amplitude also suffers from structural complexities. Methods that utilize relative amplitudes of waves, which share common ray paths, such as the spectral ratio method of Nishimura et al. (1995), the amplitude ratio method of Julian and Foulger (1996), and the relative moment tensor inversion method of Dahm (1996) appear to be better suited; among them, the linear-programming approach of Julian and Foulger (1996) has been extensively used to show the existence of many non-double-couple earthquakes in different volcanic/geothermal environments (e.g., Foulger et al., 2004; Ross et al., 1996).

An arbitrary moment tensor observed at volcanoes may be decomposed as a linear combination of those four end-member moment tensors (plus a double couple). As such a decomposition is nonunique, there can be different decompositions for a given moment tensor, and thus, there exists always nonuniqueness in interpreting moment tensor solutions in terms of actual phenomenon that is occurring at volcanoes. It is often a custom to choose the simplest possible one or to choose one that is most consistent with known source geometry of corresponding volcano (e.g., the presence of known crack-like conduit). Without such a priori information, on the other hand, one may wish to display a full moment tensor. The source-type plot introduced by Hudson et al. (1989) may be a useful tool to help interpret different moment tensors observed in different volcanic environments (e.g., Foulger et al., 2004; **Figure 10**). It displays moment tensors without regard to their orientations, and different types of

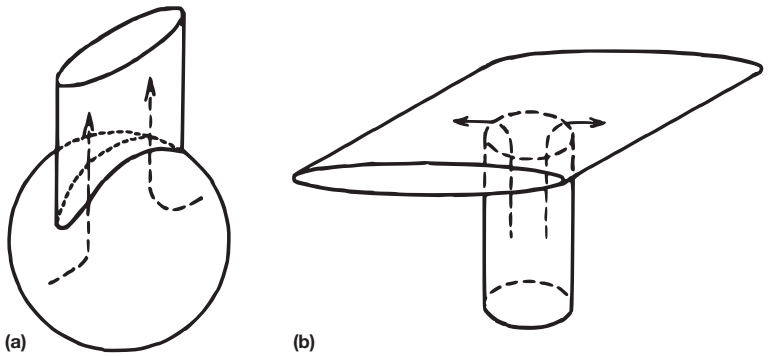


Figure 9 Geometry of selected transport models. Adapted from Chouet B (1996) New methods and future trends in seismological volcano monitoring. In: *Monitoring and Mitigation of Volcano Hazards*, pp. 23–97. New York: Springer-Verlag.

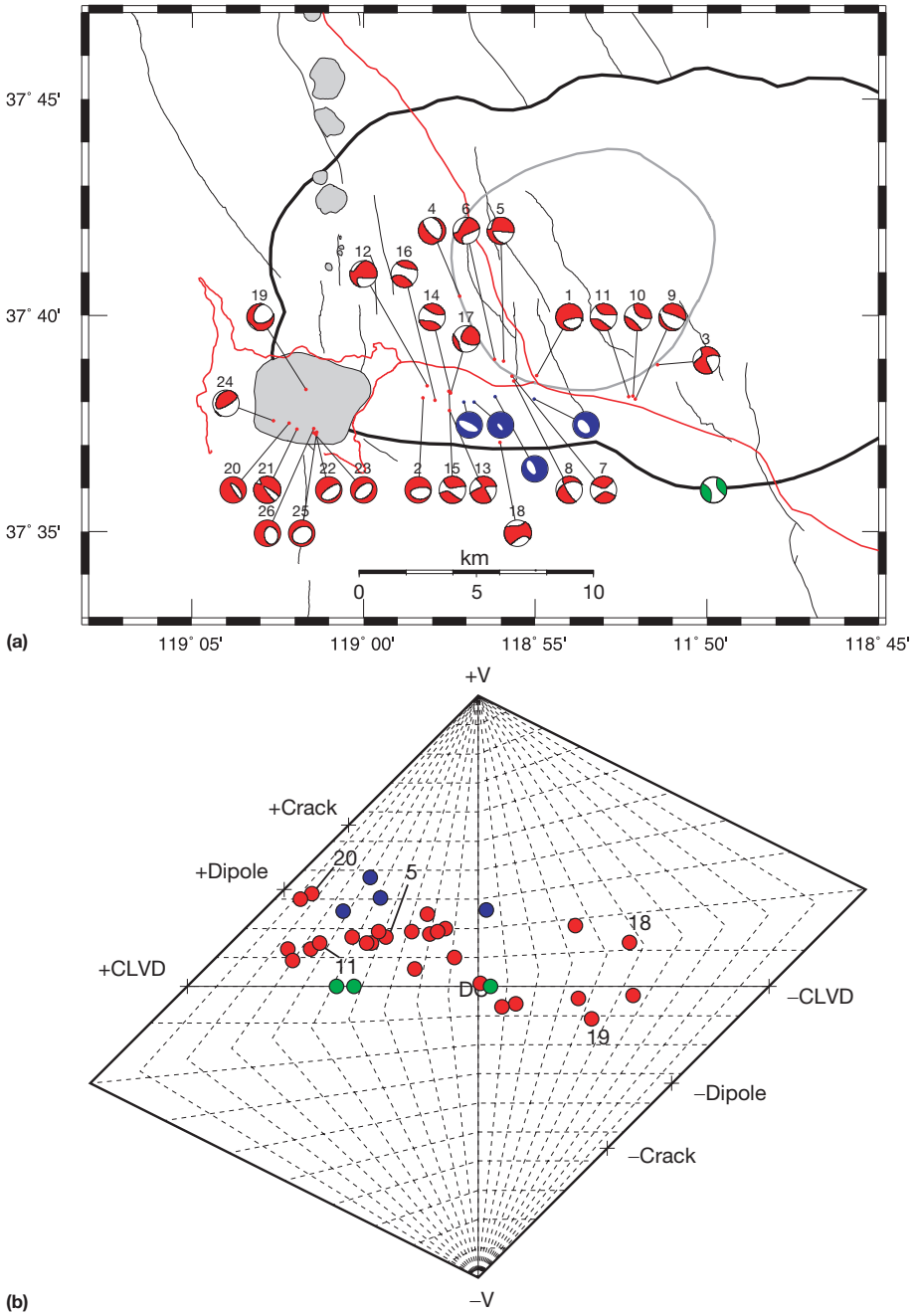


Figure 10 (a) Focal mechanism plot (upper focal hemisphere) of earthquakes in Long Valley. (b) Source-type plot of (Hudson et al., 1989) for the same earthquakes. The vertical axis measures the size of volumetric component, and deviatoric moment tensors are located on the horizontal axis (a pure double couple at the origin). Different colors indicate mechanism solutions determined by different researchers. Adapted from Foulger GR, Julian BR, Hill DP, Pitt AM, Malin PE, and Shalev E (2004) Non-double-couple microearthquakes at Long Valley caldera, California, provide evidence for hydraulic fracturing. *Journal of Volcanology and Geothermal Research* 132: 45–71.

moment tensors (different in relative magnitudes of principal moments) are projected on a rhombic plane in such a way that the area is proportional to the probability of the occurrence of different source types.

4.15.3.3 The Single Force

As described earlier, the single-force equivalent body force is due to a momentum exchange between the source volume and the rest of the Earth. Whenever a part of the Earth is detached from the rest and gains a momentum, the counterforce of this acceleration is felt by the rest of the Earth. As the detached mass eventually has to stop somehow, there must be a deceleration stage in this process, giving another counterforce with the opposite direction. The total of these two counterforces must cancel out as required from eqn [9].

The single-force equivalence of LP seismic waves excited by a landslide and explosions associated with the 1980 eruption

of Mt. St. Helens was established by Kanamori and Given (1982) and Kanamori et al. (1984). At an LP limit, a landslide may be viewed as a sliding of a box along a slope (Hasegawa and Kanamori, 1987; Kawakatsu, 1989), and its single-force equivalence may be obvious. Kawakatsu (1989) developed a centroid single-force (CSF) inversion method, an extension of the centroid moment tensor inversion method of Dziewonski et al. (1981) to include single-force components, and applied it successfully to several landslide/slump events recorded by global seismic networks. Ekström et al. (2003) applied the CSF inversion to 'glacial earthquakes' observed in southern Alaska and showed that they are represented by stick-slip, downhill sliding of a glacial ice mass, which takes about 30–60 s in duration. Kanamori et al. (1984) also suggested that an equivalent force system for a volcanic eruption can be decomposed into a vertical downward single force and an isotropic moment tensor (**Figure 11**). A single-force source may be also realized when high-viscous magma flows along a narrow conduit, as suggested by Ukawa and Ohtake (1987) to explain the deep monochromatic LF (1 Hz) event observed beneath Izu-Oshima Volcano, Japan, about 1 year prior to the 1986 eruption. Chouet et al. (2003) also suggested that a piston-like rise of a slug of gas in the conduit produced observed single-force components in the LP signals associated with explosions in Stromboli Volcano, Italy.

4.15.3.4 Waveform Analysis of Volcanic Seismic Sources

The LP part of volcanic seismic signals has been recorded by seismic networks of regional and/or global scales. One of the advantages of using such LP signals is that waveform analysis to infer excitation mechanisms can be easily performed compared to short-period seismograms. As mentioned, Kanamori and Given (1982) modeled the LP seismic signals accompanied with the eruption of the Mt. St. Helens in 1980 recorded at stations of the global seismic networks. They showed that the single-force mechanism can explain the radiation pattern of LP surface waves and attributed it to the gigantic mass movement of the volcano edifice due to a landslide. Kanamori et al. (1984) also analyzed the LP body waves generated by the same eruption and showed that vertical single forces due to the depressurization of the magma chamber explain the data. With these two single-force mechanisms, they could vividly picture what had occurred during the whole sequence of the eruption. These two studies initiated the applications of 'LP seismology' to study volcanic activities.

For LP seismic signals, it is often appropriate to treat the source originating from a point in space. Then, seismic wave field $u_n(x,t)$ may be written as

$$u_n(x,t) = G_{np,q}(x,x_0,t) * M_{pq}(t) + G_{np}(x,x_0,t) * F_p(t)$$

where $M_{pq}(t)$ and $F_p(t)$ denote a moment tensor and single-force source time function at x_0, respectively, and $G_{np}(x,x_0,t)$ and $G_{np,q}(x,x_0,t)$, respectively, represent Green's function and its spatial derivative (e.g., Aki and Richards, 2002); the asterisk denotes convolution, and the summation convention for repeated indices is assumed. Knowing the Earth structure (i.e., knowing Green's function) and having enough waveform data, it is then possible to solve for the time history of general seismic sources, $M_{pq}(t)$ and $F_p(t)$.

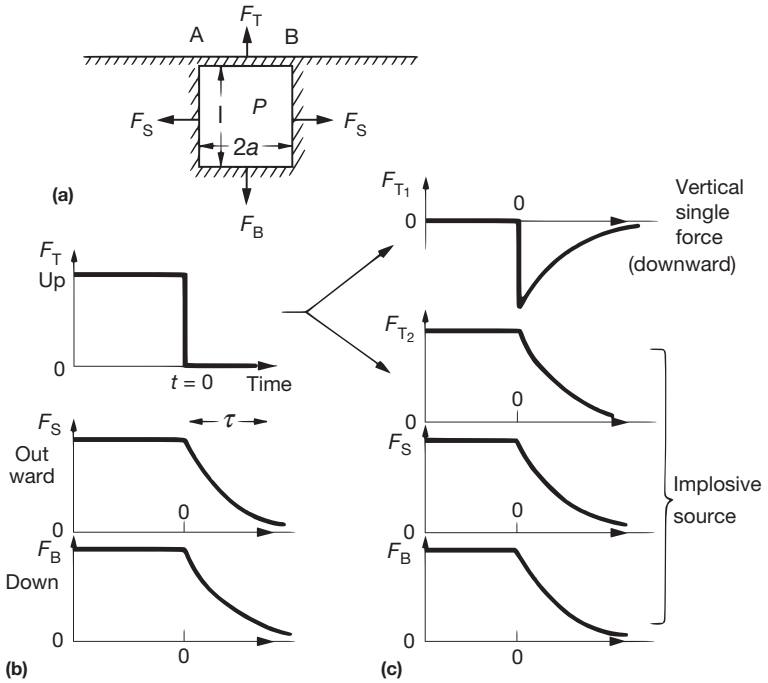


Figure 11 A force system equivalent to a volcanic eruption. (a) A pressure release model for a volcanic eruption. (b) Forces acting on the top, side, and bottom walls. (c) Decomposition of the force to a vertical single force and an implosive force. Adapted from Kanamori H, Given JW, and Lay T (1984) Analysis of seismic body waves excited by the Mount St. Helens eruption of May 18, 1980. *Journal of Geophysical Research* 89: 1856–1866.

Takeo et al. (1990) extended the waveform moment tensor inversion technique commonly used in the analysis of far-field records in earthquake seismology (e.g., Kikuchi and Kanamori, 1982) to volcano seismology by including the single-force contribution. They applied it to the records observed in a lava drain-back stage of the 1987 eruption of Izu-Oshima Volcano, Japan, and showed that the resolved vertical single forces with opposing directions are due to a rapid collapse of overloading lava into the deeper vent. Ohminato et al. (1998) applied a similar technique to near-field broadband records observed at Kilauea Volcano, USA, and showed that LP signals can be explained by a subhorizontal crack, which periodically injects magma into a larger reservoir. Chouet et al. (2003, 2005) and Ohminato (2006) further incorporated the effect of complex topography of volcanic edifice that may significantly distort seismic signals observed in the near-field by calculating Green's functions using the finite difference scheme developed by Ohminato and Chouet (1997).

The general source analysis for LP signals using moment tensor and single force has become common in seismological studies of volcanic activities. Examples of such studies include Mt. St. Helens (Kanamori and Given, 1982; Kanamori et al., 1984; Kawakatsu, 1989), Long Valley (Aki, 1984; Dreger et al., 2000; Julian, 1983; Wallace, 1985), and Kilauea, USA (Chouet et al., 2010; Eissler and Kanamori, 1987; Ohminato et al., 1998); Asama (Ohminato et al., 2006; Takeo et al., 1984), Izu-Oshima (Takeo et al., 1990), Sakurajima (Uhira and Takeo, 1994), Unzen (Uhira et al., 1994b; Yamasato et al., 1993), Ito-oki (Takeo, 1992), and Tori-shima, Japan (Kanamori et al., 1993); Hengill-Grensdalur volcanic complex (Julian et al., 1997), Bardarbunga (Nettles and Ekstrom, 1998), Iceland; Stromboli, Italy (Chouet et al., 2003, 2008); and Popocatépetl, Mexico (Chouet et al., 2005). Applications of such analysis techniques for shorter-period ($\sim$1 s) near-field records obtained at volcanoes are also becoming available (e.g., Aoyama and Takeo, 2001; Kumagai et al., 2005; Nakamichi et al., 2003; Nakano et al., 2003; Nishimura et al., 1995; Ohminato, 2006; Uhira et al., 1994a).

4.15.4 Physical Mechanisms for Volcanic Seismic Signals

The presence of ample fluid (gas, vapor, magma, or their mixtures) in volcanic edifice introduces an additional interesting class of vibration/wave phenomena to the well-studied elastic formulations. For example, the so-called crack wave, which is introduced by Chouet (1986) to explain the LF nature of volcanic events, may be understood as a part of a class of waves due to solid/fluid coupling. Julian (1994) showed that the nonlinear effect of such solid/fluid coupling due to fluid flow transient can generate a wide variety of waveforms, which have similar characteristics as volcano seismic signals. This section summarizes some such topics relevant to generation of volcano-specific seismic signals.

4.15.4.1 Slow Waves in Solid/Fluid Composite

There exists a class of waves in the two-phase system of a solid/liquid composite that travel slower than any of the sound velocities of the pure material constituting the two-phase system (i.e., the wave speed can be slower than that of the liquid phase). These waves include the tube wave (Biot, 1952), the so-called Biot's slow wave in a porous medium (Biot, 1956), the so-called crack wave (Chouet, 1986; Ferrazzini and Aki, 1987), and waves in solid/liquid alternating layers (Schoenberg and Sen, 1983) (Rayleigh wave and Stoneley wave may also be included in the same class, considering that in the high-frequency limit, all the previous slow waves have the same characteristics as an interface wave).

The simplest (and the most relevant to volcano seismology) example of slow waves may be the plane-wave propagation in a medium, which consists of a fluid layer sandwiched by two elastic half-spaces. For this problem, Krauklis (1962) first derived an analytic solution, and later, Ferrazzini and Aki (1987) reformulated the problem aiming to investigate the physical basis of the crack wave found by Chouet (1986). Here, we follow Yamamura (1997) to formulate the problem.

Following Aki and Richards (2002, p. 263), we first express a plane-wave solution propagating parallel to the fluid layer as

$$\begin{pmatrix} u \\ w \end{pmatrix} = \begin{pmatrix} r_1(z) \\ ir_2(z) \end{pmatrix} e^{i(kx-\omega t)} \qquad [18]$$

where u and w are, respectively, the displacements for x- and z-directions. The frequency domain equation of motion in x-direction for the fluid layer may then be written as

$$-\omega^2 \rho_f r_1 = -k^2 \kappa_f r_1 - k\kappa_f \frac{\partial r_2}{\partial z} \qquad [19]$$

where ρ_f and κ_f denote density and bulk modulus of the fluid, respectively. Integrating [19] over the thickness of the fluid layer (i.e., z-direction) with appropriate boundary conditions (e.g., free-slip condition at the solid/fluid interface), we get the following expression:

$$-\omega^2 \rho_f \int_0^H r_1 dz = -k^2 \kappa_f \int_0^H r_1 dz - k\kappa_f r_2(H) \qquad [20]$$

where we assume a symmetrical solution for the z-dependency (Ferrazzini and Aki, 1987) and the thickness of the fluid layer $2H$.

Equation [20] clearly demonstrates the role of the interface: without the last term $r_2(H)$, eqn [20] indicates the averaged displacement in the fluid layer traveling with its own sound velocity; that is, the term represents the way the fluid layer is coupled with the surrounding elastic half-space. If we further rewrite eqn [20] as

$$-\omega^2 \rho_f \int_0^H r_1 dz = -k^2 \kappa_f \left[1 + \frac{1}{k} \frac{r_2(H)}{\int_0^H r_1 dz} \right] \int_0^H r_1 dz \qquad [21]$$

it is seen that

$$\kappa_f^e \equiv \kappa_f \left[1 + \frac{1}{k} \frac{r_2(H)}{\int_0^H r_1 dz} \right]$$

represents the effective bulk modulus for the fluid layer; when $\kappa_f^e < \kappa_f$, that is,

$$\frac{r_2(H)}{\int_0^H r_1 dz} < 0 \qquad [22]$$

the corresponding traveling wave has a smaller wave speed relative to the sound velocity; thus, we have a slow wave. The condition [22] is satisfied if the denominator and numerator have opposite signs to each other. This makes quite good sense intuitively; when the fluid 'tries' to contract (or expand) as a passage of the wave, due to the 'elastic coupling' along the solid/fluid boundary, there exists a solution in which the solid behaves out of phase to the fluid to expand (contract), deforming the boundary in such a way that fluid 'feels' easy to contract (expand), compared to the single-phase case; that is, for a given amount of displacement in the propagation direction, the actual pressure increment of the fluid phase becomes lower than that of the pure fluid case (**Figure 12**). The net effect of this elastic coupling is to reduce the 'effective bulk modulus' (i.e., the restoring force) of the fluid, and as a result, the wave speed becomes slower than that of the pure fluid; the stronger the elastic coupling, the slower the wave speed (**Figure 13**).

This kind of the out-of-phase condition is satisfied by other slow waves such as Biot's slow wave in the porous medium

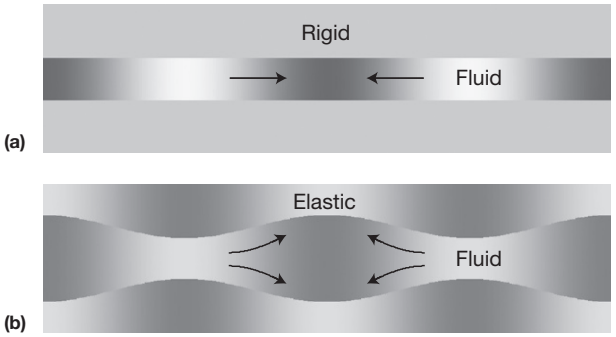


Figure 12 Conceptual image of the slow wave in solid/fluid composites. Out-of-phase motion of the solid (b) lowers the 'effective bulk modulus' of the fluid, and as a result, the wave speed becomes slower than that in the pure fluid (a). (after Yamamoto, 2005)

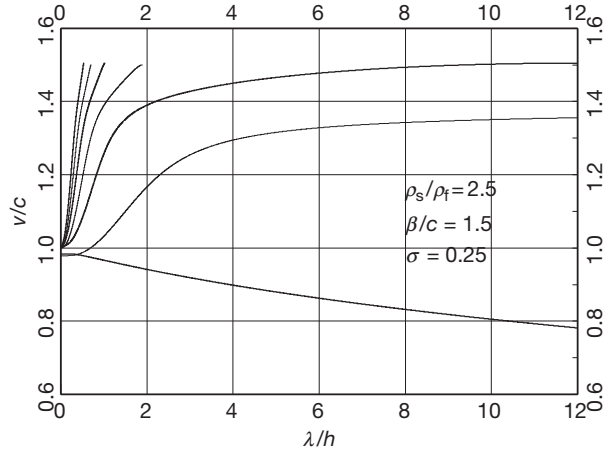


Figure 13 Phase velocity divided by the acoustic velocity of the fluid is plotted as a function of the wavelength divided by the fluid layer thickness. The fundamental branch having a phase velocity below 1 corresponds to the slow wave. Adapted from Ferrazzini V and Aki K (1987) Slow waves trapped in a fluid-filled infinite crack: Implication for volcanic tremor. *Journal of Geophysical Research* 92: 9215–9223.

(Biot, 1956) and waves in solid/fluid alternating layers (Schoenberg and Sen, 1983). Indeed, using the argument presented above, Yamamura (1997) was able to derive a similar set of equations to describe the elastic coupling for the Schoenberg's problem to those Biot's equations of motion and suggested that all these slow waves have a similar characteristic as an interface wave and that, in the short-wavelength limit, they reduce to the Stoneley wave.

4.15.4.2 Resonating Sources: A Crack

The presence of regularly spaced spectral peaks often observed in volcanic seismic signals immediately reminds us of the presence of resonators within the volcanic edifice filled with volcanic fluid. To quantitatively model the volcanic tremor observed at Kilauea, Hawaii, Aki et al. (1977) considered fluid-filled tensile cracks as the most plausible resonating sources. Using the two-dimensional finite difference method, they computed the dynamics of a series of fluid-filled cracks connected by narrow channels, which are excited by jerky extension/opening of channels due to excess fluid pressure. Although their model offers both the driving force and the geometry adequate for the magma transport beneath active volcanoes, the fluid inside the crack acts passively like a cushion so that the stress at the crack wall depends only on local displacements of the crack wall; the fluid also does not support the propagation of pressure perturbation caused by the deformation of the crack wall. To assess active roles of the fluid, Chouet and Julian (1985) extended Aki et al.'s model and investigated dynamic interaction between fluid and elastic solid inside and outside a two-dimensional crack by solving equations for the elastic solid and the fluid simultaneously. The model was further extended to three dimensions by Chouet (1986), and Yamamoto and Kawakatsu (2008) also formulated the two-dimensional problem using the boundary integral method.

Chouet's three-dimensional model consists of a single isolated fluid-filled crack in an infinite elastic solid body, and no external mass transfer into and/or out of the crack is taken into consideration. In the model, the thickness of the crack d is assumed to be much smaller than the wavelength of interest, and thus, the motion of the fluid inside the crack is treated as two-dimensional. Assuming Poisson's ratio to be 0.25 (i.e., $\lambda = \mu$) and taking L and L/α as the characteristic length and time, where L and α are the length of the crack and the P wave velocity in the solid, the response of the crack is calculated by solving the following dimensionless equations simultaneously with boundary conditions at the crack surface and perimeters:

$$\begin{aligned}
\text{Solid}: \quad & \frac{\partial \tilde{v}_i}{\partial \tilde{t}} = \frac{1}{3}\frac{\partial \tilde{\tau}_{ik}}{\partial \tilde{x}_k} \\
& \frac{\partial \tilde{\tau}_{ij}}{\partial \tilde{t}} = \frac{\partial \tilde{v}_k}{\partial \tilde{x}_k}\,\delta_{ij} + \frac{\partial \tilde{v}_i}{\partial \tilde{x}_j} + \frac{\partial \tilde{v}_j}{\partial \tilde{x}_i} \\
\text{Fluid}: \quad & \frac{\partial \tilde{U}_l}{\partial \tilde{t}} = -\frac{1}{3}\,\frac{\rho}{\rho_f}\frac{\partial \tilde{P}}{\partial \tilde{x}_l} \\
& \frac{\partial \tilde{P}}{\partial \tilde{t}} = -\frac{\kappa_f}{\mu}\frac{\partial \tilde{U}_m}{\partial \tilde{x}_m} - 2\,\frac{\kappa_f}{\mu}\frac{L}{d}\,\tilde{v}_z \\
& (\tilde{x} = x/L,\ \tilde{t} = t\,\alpha/L)
\end{aligned} \qquad [23]$$

In these equations, the quantities with tilde represent dimensionless ones, and $\tilde{v}$ and $\tilde{\tau}$ are the velocity and the stress

of the solid; $\tilde{U}$ and $\tilde{P}$ are the velocity and the pressure of the fluid, ρ and ρ_f are the density of the solid and the fluid, and μ and κ_f are the rigidity of the solid and the bulk modulus of the fluid, respectively. As shown in eqn [23], the behavior of fluid-filled crack critically depends on a nondimensional parameter $c = \kappa_f/\mu \cdot L/d$, which was first introduced by Aki et al. (1977) and named as 'crack stiffness' by Chouet (1986).

Synthetic seismograms calculated by the fluid-filled model show strong similarities to observed volcanic signals in terms of both peaked spectra and long-lasting oscillations and demonstrated the importance of active participation of fluids in source dynamics generating volcanic seismic signals. These studies include the interpretations of volcanic signals observed at Redoubt, Alaska (Chouet et al., 1994); Galeras, Colombia (Gil Cruz and Chouet, 1997; **Figure 14**), Kusatsu-Shirane, Japan (Kumagai et al., 2002; Nakano et al., 1998), Kilauea, Hawaii (Kumagai et al., 2005), and other volcanoes.

Kumagai and Chouet (1999, 2000) and Morrissey and Chouet (2001) further studied the dependencies of frequencies and attenuations of crack-resonant oscillations on the properties of fluid in more detail using various models of fluid–gas mixtures and fluid–particle mixtures and demonstrated that the fluid properties and compositions can be estimated from frequencies and attenuations of observed signals. These results are used to assess the change of magmatic and hydrothermal system beneath the volcano at, for example, Kusatsu-Shirane, Japan (Kumagai et al., 2002; **Figure 15**), Tungurahua, Ecuador (Molina et al., 2004), and Etna, Italy (Lokmer et al., 2008).

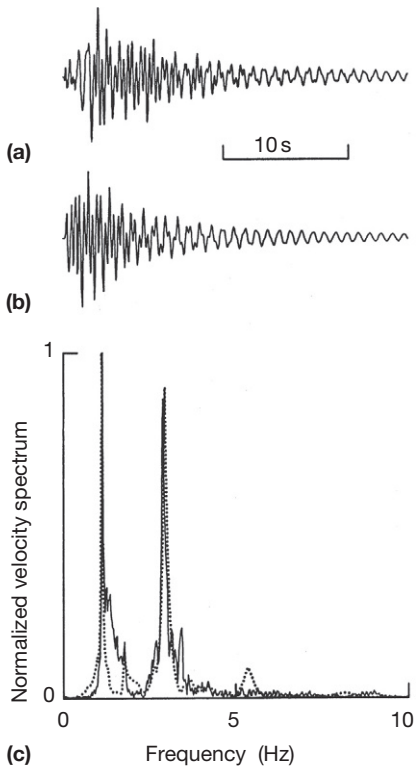


Figure 14 Example of the application of the fluid-filled crack model: (a) Vertical velocity seismogram observed at Galeras Volcano, Colombia. (b) Synthetic waveform obtained from the fluid-filled crack model. (c) Spectrum of observed data (thin line) and synthetic data (dotted line). Adapted from Chouet B (1996) Long-period volcano seismicity: Its source and use in eruption forecasting. *Nature* 380: 309–316.

4.15.4.3 Other Resonating Sources

Although the presence of regularly spaced spectral peaks may not straightforwardly warrant the presence of a resonator (Julian, 1994), it seems indeed natural to think of such presence like a crack as a dike, a cylinder as a conduit, and a sphere as a magma chamber, as they are all fundamental constituents of a volcano.

Resonators as models of the origins of volcanic seismic signals besides a crack have been suggested for a cylinder (Chouet, 1985; Ferrick et al., 1982), and a sphere (Kubotera, 1974; Sassa, 1935). Crosson and Bame (1985) obtained an impulsive response of a spherical magma chamber with a spherical cavity inside. Fujita et al. (1995) solved eigenoscillation of a fluid sphere embedded in an infinite elastic medium. Fujita and Ida (2003) conducted a systematic survey of the geometric effects on the eigenoscillations and showed that the ratios of higher modes to the fundamental mode frequencies can be indicators of the geometry of the resonator. They also found a class of modes with a low attenuation named 'low-attenuation mode' (LAM), which share the similar out-of-phase characteristics of the slow waves discussed earlier, and suggested that LAM may be the origin of long-lasting LF oscillation of volcanic seismic signals.

The complicated nature of vibrations of cylindrical conduits has been numerically studied (e.g., Jousset et al., 2003; Neuberg et al., 2000; Nishimura and Chouet, 2003). Neuberg et al. (2000) introduced a depth-dependent seismic velocity model to account for the varying gas content in the magma and showed that a pressure change in the conduit was the most likely candidate for the physical process causing the peak-spectral shift often observed in many of volcanoes (**Figure 4**). Jousset et al. (2003) also showed that the endpoints of the conduit can act as secondary sources of volcanic seismic signals.

The excitation mechanism of these resonators is, on the other hand, not well understood. The proposed models include a jerky extension of the crack tip (Aki et al., 1977), acoustic emissions from collapsing bubbles (Chouet, 1992), rapid discharge of fluids (Ferrick et al., 1982), and a sudden pressure drop caused by unsteady choked flow (Morrissey and Chouet, 1997). Nakano et al. (1998) developed a method to extract the effective excitation function from seismic records based on an inhomogeneous autoregressive model of a linear dynamic system. Nakano et al. (2003) further extended the method to deal with the nonorthogonal nature of eigenfunctions of a resonator embedded in a rock matrix by using the biorthogonal eigenfunction expansion method of Yamamura and Kawakatsu (1998); the resolved effective excitation functions of the LF events in Kusatsu-Shirane Volcano, Japan, are then modeled by the waveform inversion method of Ohminato et al. (1998) to obtain the moment tensor and single-force source time function, which are interpreted as a result of repeated activation of a subhorizontal crack located beneath the summit crater lake.

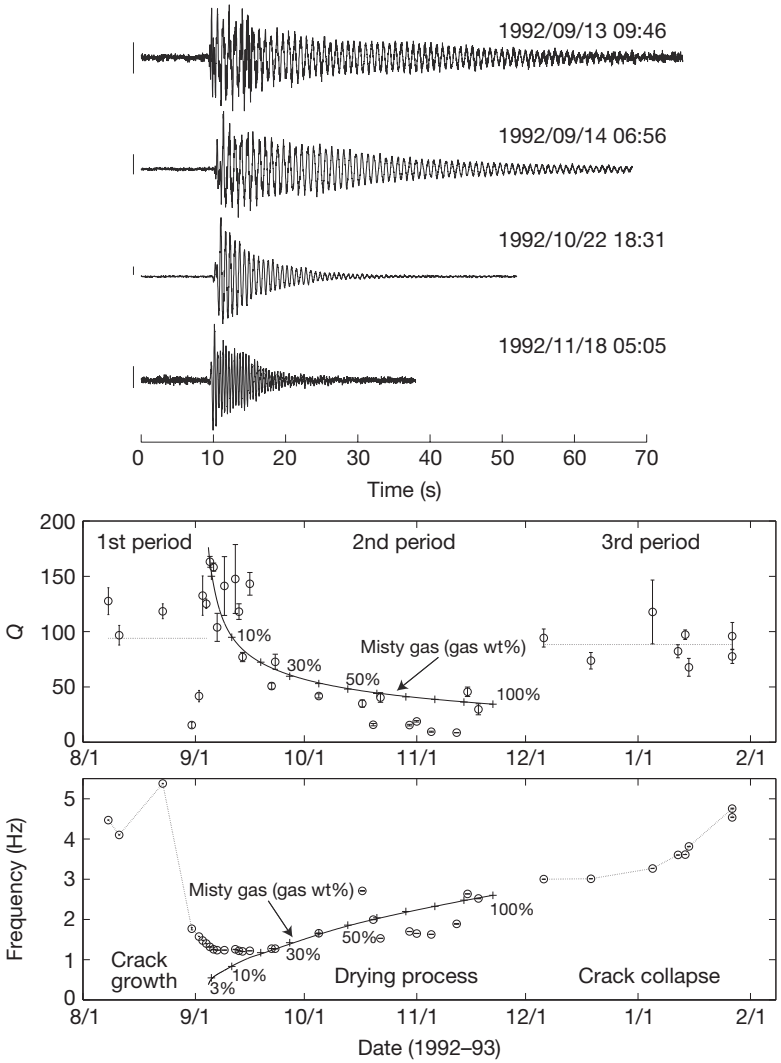


Figure 15 Temporal change of waveforms observed at Kusatsu-Shirane Volcano and the fluid properties estimated from the frequency and the attenuation of signals. Adapted from Kumagai H, Chouet B, and Nakano M (2002) Temporal evolution of a hydrothermal system in Kusatsu-Shirane Volcano, Japan, inferred from the complex frequencies of long-period events. *Journal of Geophysical Research* 107: ESE 1–ESE 9.

4.15.4.4 Flow-Induced Oscillation

Julian (1994) proposed nonlinear flow-induced oscillations in channels transporting volcanic fluid as a possible generating mechanism for volcanic seismic signals, particularly for sustained volcanic tremors. The mechanism may have analogs such as the vibration of vocal chords and musical wind instruments, and the periodic behavior, often characteristic of volcanic seismic signals, occurs without the presence of a resonator.

Figure 16 shows the simplified lumped-parameter model of the flow of a viscous fluid through a constricted channel with elastic walls. The channel walls are modeled as masses connected by a spring and dashpot to represent the effects of the elastic stiffness and radiation loss. The mechanism may be summarized as follows: an increase in a flow speed leads to a decrease in fluid pressure by the Bernoulli effect; as a result, the walls narrow the channel to constrict the flow, causing a pressure increase and forcing the channel open again; this again decreases the pressure and increases the flow speed. This positive feedback mechanism sustains the oscillation.

Julian (1994) derived a third-order nonlinear system of ordinary differential equations that describe the complete behavior of the system and may be solved numerically. The system without radiation damping exhibits two different types of behavior: oscillations that grow and approach a stable tremor-like limit cycle and transient oscillations that decay, leading to a steady flow. Inclusion of the damping further complicates the behavior, as shown in Figure 17. By increasing the upstream pressure, the system experiences the so-called period doubling, that is, an appearance of a subharmonic with a frequency half that of the fundamental frequency; further increase with successive period doubling eventually leads to chaotic behaviors. Julian suggests that these features resemble behavior of some of the volcanic tremors and the deep LF event observed in Hawaii (Aki and Koyanagi, 1981).

Balmforth et al. (2005) extended Julian's treatment to include the dynamic behavior of the fluid and the elasticity

of the surrounding walls and considered the excitation of propagating waves. Their results confirm that Julian's mechanism is plausible, but the physical condition required for the instability is different. They conclude that magma itself is unlikely to generate flow-induced oscillations but that the rapid flow through fractured rock of low-viscosity fluids exsolved from magma may. Similar flow-induced oscillations are also studied by Ida (1996), Corona-Romero et al. (2012), and Dunham and Ogden (2012).

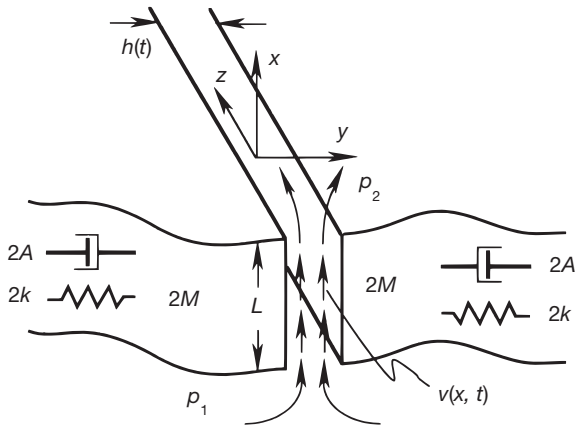


Figure 16 Lumped-parameter model of the generation of volcanic tremor. Viscous, incompressible fluid flows in x-direction from upstream (bottom) to downstream (top) reservoir through a channel with imperfectly elastic walls. Adapted from Julian BR (1994) Volcanic tremor: Nonlinear excitation by fluid flow. *Journal of Geophysical Research* 99: 11859–11878.

4.15.4.4.1 *In fluid/gas two-phase system*

The flow-induced oscillation described earlier assumes flows of a single-phase fluid. Volcanic fluids such as magma often contain ample gases, especially at a shallow depth, and it may be more reasonable to consider a flow of two-phase fluids consisting of liquid and gas phases rather than single-phase fluids as a source of volcanic seismic signals. Such a flow of liquid/gas two-phase fluid is also seen in, for example, the evaporation pipes in boiler systems, nuclear reactor cooling systems, and chemical industry equipments, and many studies have been conducted extensively in the field of engineering. These studies revealed the existence of flow-induced oscillatory phenomena, which is not observed in a flow of single-phase fluid, caused by complex interactions between two phases.

Iwamura and Kaneshima (2005) considered one of these flow-instability phenomena called 'density wave oscillation,' which was originally studied in the nuclear reactor engineering, as a model of volcanic seismic signals, and evaluated the properties of generated seismic waves through a series of numerical experiments. The density wave oscillation and limit cycle of the system was also studied by Fujita (2008) and applied to the analysis of the banded tremor at Miyake-jima Volcano. Another kind of flow-induced oscillation in fluid/gas two-phase flows called 'pressure drop oscillation' was also invoked to interpret the cyclic tilt change observed at Miyake-jima Volcano by Fujita et al. (2004).

4.15.4.5 Bubble Dynamics

Gases dissolved in magma are one of the main driving forces to cause various volcanic phenomena. The presence of gas bubbles in magma decreases the density of magma and drives the

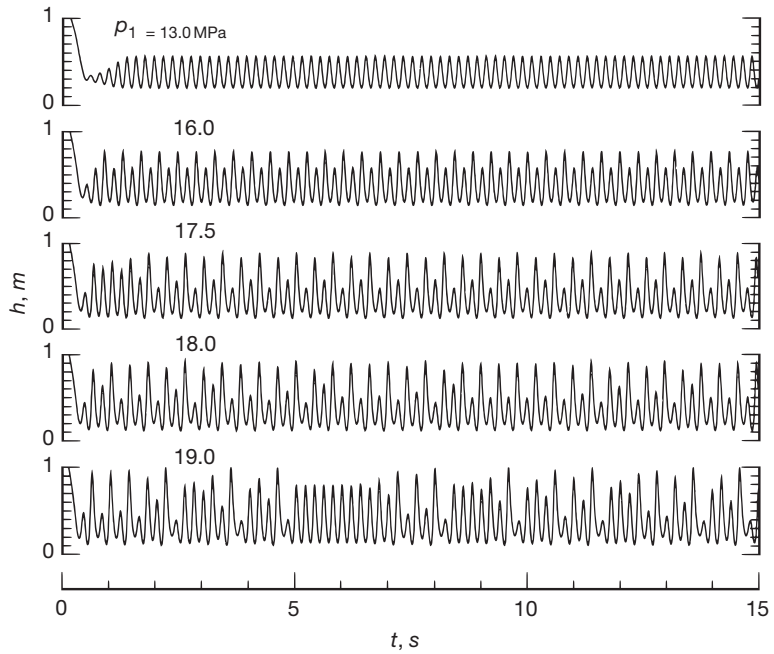


Figure 17 Synthetic tremor time series (channel thickness h vs. time) for different values of the upstream pressure p_1. Behavior changes from a simple limit cycle (top) to chaos (bottom), with some period doubling in between. Adapted from Julian BR (1994) Volcanic tremor: Nonlinear excitation by fluid flow. *Journal of Geophysical Research* 99: 11859–11878.

magma ascent from the deep part of volcanoes (e.g., Wilson and Head, 1981). At shallower depths, the relatively high compressibility of bubbles contributes to the generation of oscillatory phenomena specific to volcanoes (e.g., Chouet, 1996a; Kieffer, 1977), and the transformation and the fragmentation of bubbles affect the behavior of volcanic eruptions (e.g., Alidibirov and Dingwell, 1996; Ichihara et al., 2002).

In relation to volcano seismology, there are some aspects where bubbles play important roles. As described earlier, volcanic seismic signals are often characterized by long-lasting LF oscillations with sharp spectral peaks (**Figure 1**), which are reminiscent of oscillations of some resonator filled with material with a slow acoustic velocity. Since the existence of bubbles in magma drastically changes the compressibility of magma, as pointed out by Kieffer (1977), these characteristics of volcanic seismic signals have been partly attributed to the presence of bubbles in magma. The rapid temporal changes of the observed spectral peaks (e.g., Benoit and McNutt, 1997; Neuberg et al., 2000) have also been considered to reflect the change in the state of gas bubbles in volcanic conduits.

Bubbles in magma also act as an active source of seismic and acoustic signals: The bursting and the collapsing of bubbles at the lava surface preceding and/or accompanying volcanic eruptions generate acoustic pressure signals, which are sometimes accompanied with seismic waves (e.g., Ripepe et al., 1996; Vergniolle et al., 1996); the collapse of small bubbles in hydrothermal fluids generates seismic tremors (e.g., Leet, 1988); the movement of large-scale bubbles (slugs) in magma conduit causes an exchange of their momentum with the surrounding rocks and generates seismic signals that show dominant single-force components (e.g., Chouet et al., 2003, 2005, 2008, 2010; Ohminato et al., 1998; Ripepe and Gordeev, 1999); the forced oscillations of bubbles by strain waves can rapidly pump dissolved volatiles into bubbles by a mechanism called 'rectified diffusion' and the resultant pressure increase may trigger volcanic activities (e.g., Brodsky et al., 1998; Ichihara and Brodsky, 2006).

In spite of the difference in their generation mechanisms and resultant phenomena, all of these phenomena are the manifestations of passive and active processes in which the thermochemical energy and/or the gravitational energy of the volcanic fluids is transduced into seismic/acoustic energy. The quantitative understanding of the behavior of bubbles, thus, is a critical key to elucidate the dynamics of volcanic phenomena.

4.15.5 Observation and Analysis Aspects

In addition to the conventional monitoring of active volcanoes using short-period seismometers and geodetic instruments, the digital era has brought new powerful components to seismometry of active volcanoes; broadband and array seismometries may be two of the most notable developments.

4.15.5.1 Broadband Seismometry

The first broadband seismic observation at an active volcano was conducted by Sassa (1935), who installed Wiechert horizontal-component seismographs (pendulum period $T_0 = 10.0$ s) and vertical ones ($T_0 = 4.6$ s), Gallitzin seismographs ($T_0 = 8.0$ s), and short-period seismographs ($T_0 = 0.55$ s) at Aso Volcano, Japan. Summarizing the subsequent observations in a period of several years, Sassa classified volcanic seismic signals of Aso into 'eruption earthquakes' and four different kinds of tremors: periods of 0.2, 0.4–0.6, 0.8–1.5, and 3.5–8.0 s. The observation of these tremors in the wide frequency band ranging from 0.2 to 8.0 s clearly demonstrates the efficacy of broadband seismometry at active volcanoes. Since Sassa's pioneering work, until 1990s (when it has become much easier to install portable broadband instruments), there have not been many published reports on attempts to observe LP volcanic seismic signals except for a few cases. For example, Seidl et al. (1981) installed broadband seismometers at Etna, Italy, and observed 4–5 s period signals, which they suggested as one of the fundamental peaks of the Etna tremors.

Examples of deployments of the current-generation broadband seismometers are now plentiful, for example, Sakurajima (Kawakatsu et al., 1992, 1994), Unzen (Uhira et al., 1994b; Yamasato et al., 1993), Aso (Kaneshima et al., 1996; Kawakatsu et al., 2000; Legrand et al., 2000; Yamamoto et al., 1999), Satsuma-Iwojima (Ohminato, 2006; Ohminato and Ereditato, 1997), Iwate (Nishimura et al., 2000), Miyake-jima (Fujita and Ida, 2003; Kobayashi et al., 2009; Kumagai et al., 2001), Usu (Yamamoto et al., 2002), Bandai (Nishimura et al., 2003), Hachijo-jima (Kumagai, 2006), and Asama, Japan (Ohminato et al., 2006); Stromboli, Italy (Chouet et al., 2003; Dreier et al., 1994; Falsaperla et al., 1994; Neuberg et al., 1994); Semeru (Hellweg et al., 1994) and Merapi, Indonesia (Hidayat et al., 2000, 2002); Arenal, Costa Rica (Hagerty et al., 2000); Kilauea (Almendros et al., 2002a; Chouet and Dawson, 2011; Chouet et al., 2010; Dawson et al., 1998; Ohminato et al., 1998), Long Valley (Hill and Prejean, 2005; Hill et al., 2002), and St. Helens, USA (Waite et al., 2008); Erebus, Antarctica (Aster et al., 2003, 2008; Rowe et al., 1998, 2000); Popocatépetl, Mexico (Arciniega-Ceballos et al., 1999, 2003, 2008; Chouet et al., 2005); and Cotopaxi, Ecuador (Molina et al., 2008).

Varieties of LP volcanic seismic signals have been observed at different volcanoes. Since the broadband seismometers are also sensitive to tilt motion, which can be analytically described (e.g., Graizer, 2006; Rodgers, 1968), they have been also used to analyze the deformation due to movements of magma in shallow volcanic edifice (e.g., Aoyama and Oshima, 2007; Wielandt and Forbriger, 1999).

Besides the advantages already mentioned, the LP seismic wavefield observed with a broadband seismic network may be used to directly monitor the activity of a volcano. Dawson et al. (2004) implemented the radial semblance method of Kawakatsu et al. (2000, originally called 'waveform semblance') for the near real-time monitoring of the activities at the shallow magmatic conduit in Kilauea Volcano, Hawaii (**Figure 18**). Similarly, the moment tensor and single-force source real-time monitoring should be possible using a grid-based method as suggested by Kawakatsu (1998) and implemented at the Earthquake Research Institute of the University of Tokyo for the real-time seismicity monitoring (Tsuruoka et al., 2009; http://wwweic.eri.u-tokyo.ac.jp/GRiD_MT/) and as realized at Stromboli Volcano, Italy (Auger et al., 2006).

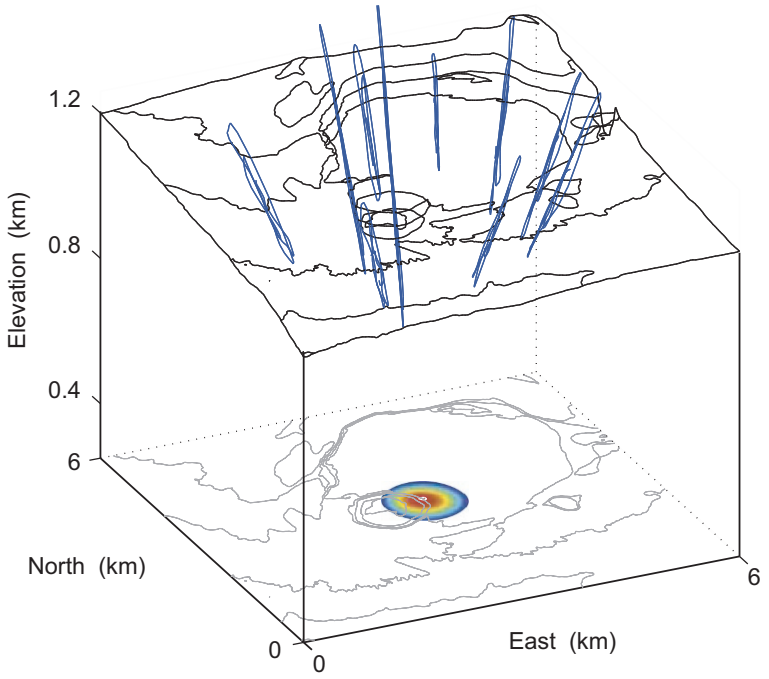


Figure 18 Particle motions recorded by the network during a 1 min long window of long-period signal. The area with bright colors with high semblance values indicates the source region. Adapted from Dawson P, Whilldin D, and Chouet B (2004) Application of near real-time radial semblance to locate the shallow magmatic conduit at Kilauea Volcano, Hawaii. *Geophysical Research Letters* 31: L21606.

Kawakatsu and Montagner (2008) recently clarified the connection between these grid-based source inversion approaches and the adjoint or time-reversal source imaging approach, by showing that the adjoint method is an approximation to the inversion problem.

4.15.5.2 Array Analysis

To locate the source of volcanic signals, various kinds of methods have been applied to the analyses of observed data. Classical hypocenter determination methods using phase arrivals have been applied to signals observed around active volcanoes and used to predict the eruptions (e.g., Scarpa and Tilling, 1996). More advanced methods such as relative hypocenter determination using the cross-spectral and the double difference have also been applied to volcanic earthquakes and have revealed vivid images of volcanic activities like the transportation of magma beneath active volcanoes (e.g., Hayashi and Morita, 2003; Morita et al., 2006).

Nevertheless, there still exist some specific issues to be explored in the analyses of volcanic signals due to complex source processes and complicated wave propagations in inhomogeneous structures. Lack of clear phases and relatively low dominant frequencies of, for example, LF events and volcanic tremors make the direct application of the conventional hypocenter determination means rather difficult. In the past decade, these difficulties have been partly overcome by the use of seismic arrays. In array observations, seismometers are placed at short distance intervals (a few tens to hundreds of meters), and thus, the wave properties of the incident waves can be examined by extracting coherent portion of the signals using the array data. Recent advance in portable instrumentation enables us to observe volcanic signals close to their sources with dense arrays and sheds new light on the understanding of the processes through the precise determination of location and distribution of volcano seismic sources. For example, the multiple signal classification technique (Goldstein and Archuleta, 1991; Schmidt, 1986) has been applied to the analyses of the array data observed at Kilauea, Hawaii (e.g., Almendros et al., 2002a,b; Goldstein and Chouet, 1994; **Figure 19**), and at Stromboli, Italy (e.g., Chouet et al., 1997), and revealed magmatic and hydrothermal activities beneath the volcanoes. Other array analysis techniques have been also applied to constrain the source locations and properties of volcanic signals (e.g., Gordeev et al., 1990; Ibanez et al., 2000).

Seismic arrays have also been used to constrain the shallow structure of volcanoes that is indispensable for the quantitative volcano seismology. Here, the wavefield observed by a seismic array is treated as a stationary stochastic one in time and space, and the statistical correlation method (i.e., spatial autocorrelation (SPAC) method) originally proposed by Aki (1957) is applied to the data to determine the wave properties such as wave types and phase velocities (e.g., Chouet et al., 1998; De Luca et al., 1997; Ferrazzini et al., 1991; Metaxian and Lasage, 1997; Saccorotti et al., 2001a).

4.15.6 Models for Volcanic Seismic Signals

Although the basic governing physics, some of which have been discussed earlier, are common, and the methods of data acquisition and analyses employed may be the same, actual phenomena that we observed at volcanoes are different from one volcano to another. To appreciate the diversity, some of

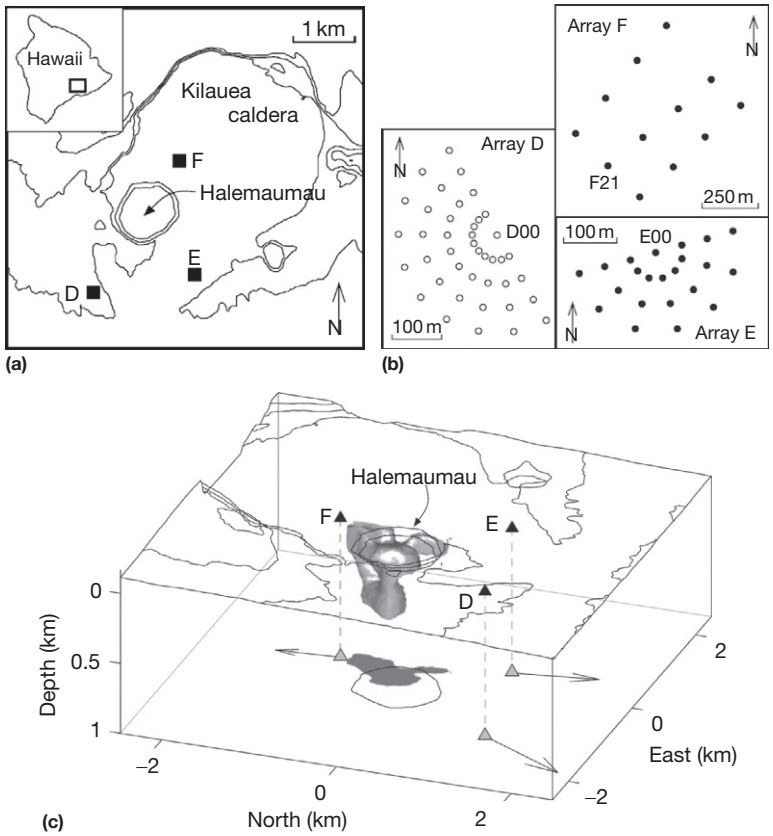


Figure 19 Example of array observation. (a and b) Configuration of the array observation at Kilauea, Hawaii. (c) Source location of the seismic signals obtained by array analyses. The arrows represent the apparent slowness vectors determined by frequency-slowness analyses, and the gray region is the estimated source region. Adapted from Almendros J, Chouet B, Dawson P, and Huber C (2002) Mapping the sources of the seismic wavefield at Kilauea Volcano, Hawaii, using data recorded on multiple seismic antennas. *Bulletin of Seismological Society of America* 92: 2333–2351.

the well-studied volcanoes are reviewed. Since the names of volcanic seismic events sometimes differ from one volcano to another due to the difference in conventions as described in 'terminology,' in this section, we keep common terminology locally used at each volcano and not necessarily follow the conventional classification of periods in earthquake seismology.

4.15.6.1 Aso

Aso Volcano is one of the most active volcanoes in Japan. It has erupted in Strombolian style repeatedly with intervals of 5–10 years, and recent activities take place at the youngest central cone, which is composed of seven craters aligned northwest–southeast direction with a length of 1 km. The current eruptive activities have occurred at the first crater located at the northernmost of the chain of the craters.

Among the volcanic tremors observed by Sassa (1935), the most remarkable one is the volcanic microtremor of the second kind for its extraordinary LP (7.5 s) and its highly repetitive occurrence. Observations using modern broadband seismometers revealed that the fundamental period of the long-period tremor (hereafter LPT) is 15 s (Kaneshima et al., 1996; Kawakatsu et al., 2000) and that the second-kind tremor of Sassa is the higher mode of LPTs. The characteristics of LPTs are summarized as follows: (1) they are continually emitted from the volcano regardless of surface activity; (2) spectra show several common spectral peaks, which align with almost equal spacing (15, 7.5, 5, and 4 s); (3) decay is fairly fast and the duration is only a few cycles; and (4) they are often accompanied with short-period tremors.

Kaneshima et al. (1996) and Kawakatsu et al. (2000) deployed a broadband seismic network in 1994 and analyzed the waveforms of the LPTs and those associated with phreatic eruptions, which ejected composite of mud and water (**Figure 6**). They concluded that the sources of both phenomena are located a few hundred meters southwest of the first crater and at a depth of 1–1.5 km below ground surface. They further performed a source mechanism inversion of LPTs and obtained a volumetric source mechanism, which can be decomposed into isotropic and vertical tensile crack components (Legrand et al., 2000). The detail of the tensile crack component was further studied by Yamamoto et al. (1999) using the data obtained by a dense broadband observation. Analyzing the spatial distribution of the signal amplitudes of LPTs (**Figure 20**), they revealed the existence of a crack-like conduit whose strike and width are almost the same as those of the chain of craters; the crack extends nearly vertically from a depth of 300–400 m below the surface to a depth of about 2.5 km, as illustrated in **Figure 21**. This observation revealed that the chain of craters is simply a surface expression of a buried crack-like conduit.

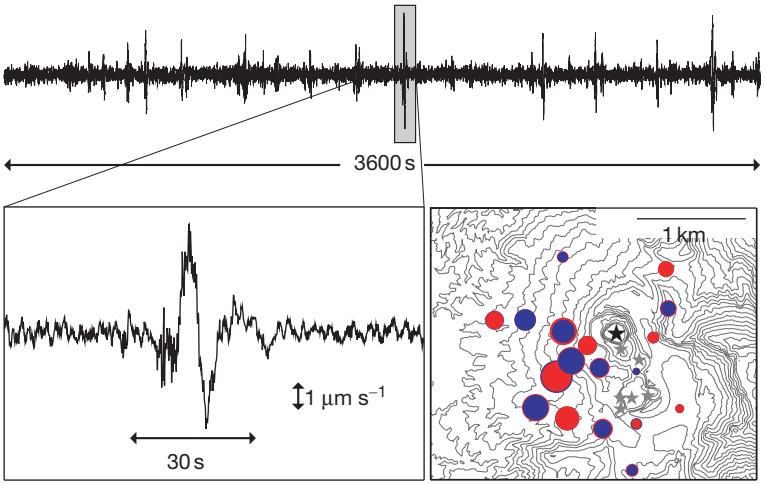


Figure 20 (a) A vertical-component broadband velocity seismogram observed at a distance of 1.5 km. Spike-like signals are LPTs. (b) Close-up view of an LPT. (c) Amplitude variation of LPTs (after Yamamoto, 2005).

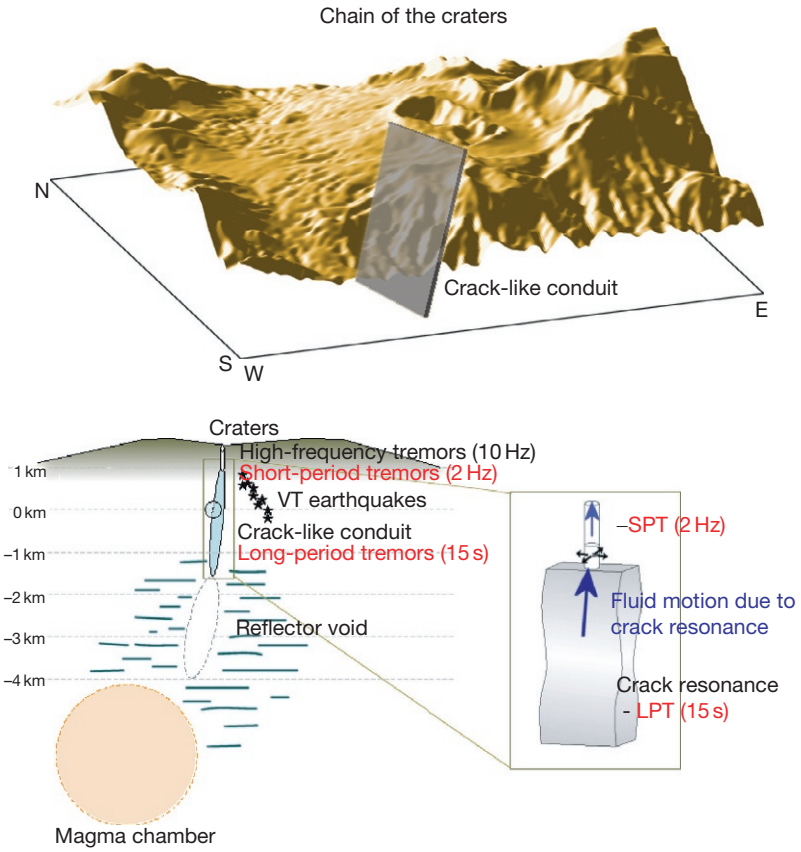


Figure 21 Image of crack-like conduit model. The crack is almost parallel to the chain of craters and the extension of the crack meets the active fumarole at the surface (after Yamamoto, 2005).

The physical properties of the crack-like conduit were studied by Yamamoto (2005), who applied a boundary integral method to the oscillation of a fluid-filled crack with slits at the edges, which allow the fluid to escape, as an extension of Chouet's closed crack model (Chouet, 1986). He demonstrated that the spectral characteristics of LPTs (mode frequencies, spacing, and attenuation) can be explained by the oscillation of a 25 m-thick crack-like conduit filled with the gas–ash mixture. The existence of a crack-like conduit is also supported by other studies such as a reflection study by Tsutsui and Sudo (2004), and the crack-like conduit is considered as a subsurface path connecting a postulated magma chamber at a depth of around 5 km (Sudo and Kong, 2001) and the surface craters. The nature of other short-period seismic signals

with a period of about 0.5 s and a period of around 0.4–0.1 s has been also studied with modern digital data obtained by dense observations using short-period seismometers (Takagi et al., 2006; Yamamoto, 2005).

Figure 21 schematically summarizes the system beneath Aso Volcano that has been revealed by seismological analyses. Such a line of volcanic conduit systems connecting the magma chamber and the ground surface has not been detected at any other active volcano in the world. Considering that the various volcanic signals are manifestations of dynamic interactions between volcanic fluids and volcanic edifice in the conduit system, Aso Volcano appears to be one of the best fields to study the dynamic behavior of the volcanic fluid system beneath active volcanoes.

4.15.6.2 Kilauea, Hawaii

The Hawaiian volcanoes have been produced by the hot spot, which is presently beneath the Hawaii islands. The volcanoes primarily erupt basaltic magma, and the relatively low viscosity of the magma yields gentle slope of the edifices. Kilauea is the youngest and southeasternmost volcano on the Hawaii islands, and it is located at the splitting point between the Southwest Rift Zone and the East Rift Zone. Owing to the high level of activities and easy access to it, Kilauea may be the best-studied volcano in the world. For instance, a cylindrical cloud of seismicity beneath the volcano, which could plausibly reflect a magma movement in the crust, was detected (e.g., Klein and Koyanagi, 1989; Wolfe et al., 2003; **Figure 22**), and its relation to the mantle fault zone caused by plate flexure beneath Hawaii has been studied (e.g., Pritchard et al., 2007).

The magma transport system at shallower depths has been also studied using broadband seismic networks and short-period seismic arrays. Performing the moment tensor inversions of broadband data, Ohminato et al. (1998) revealed that the source of the VLP (5–10 s) signals associated with a magmatic activity in 1996 is a crack-like magma pathway at a depth of 1 km, which acts like a buffer for surging magma from beneath the summit of Kilauea to the East Rift Zone. Their analyses suggest that $\sim$1000–4000 m^3 of magma is injected into the crack during a slow (1–3 min) accumulation phase, and it is ejected from the crack toward the East Rift Zone during a rapid (5–10 s) deflation phase, which together produce a sawtooth-like displacement signal shown in **Figure 23**. Dawson et al. (1998) also analyzed VLP signals observed in 1997 and concluded that the source location is almost the same as that determined by Ohminato et al. (1998), although the source mechanism is different.

The source of LP (1–2 s) signals, on the other hand, is determined using seismic array data (e.g., Almendros et al., 2001; Saccorotti et al., 2001b), and the hypocenter of LP signals is located directly above the source of VLP signals. Kumagai et al. (2005) analyzed the source mechanism of LP signals and imaged an expansion and contraction of a nearly horizontal crack through the moment tensor inversions. Based on the fluid-filled crack model developed by Chouet (1986), they further suggest that LP signals represent a resonance of a hydrothermal crack containing bubbly water and/or steam and that the heat from the underlying magma conduit (VLP source) may cause the pressurization of hydrothermal fluid in the crack and trigger these LP signals. Other geophysical observations, such as a positive correlation between summit SO_2 emissions and the shallow seismicity, which was pointed out by Sutton et al. (2001), also suggest the interaction between magmatic and hydrothermal systems at the shallow part of the volcano.

These results indicate that the detailed quantitative analyses of volcanic seismic signals are critically important for a better understanding of volcanic processes and such analyses may be crucial steps toward eruption prediction and the assessment of volcanic hazards.

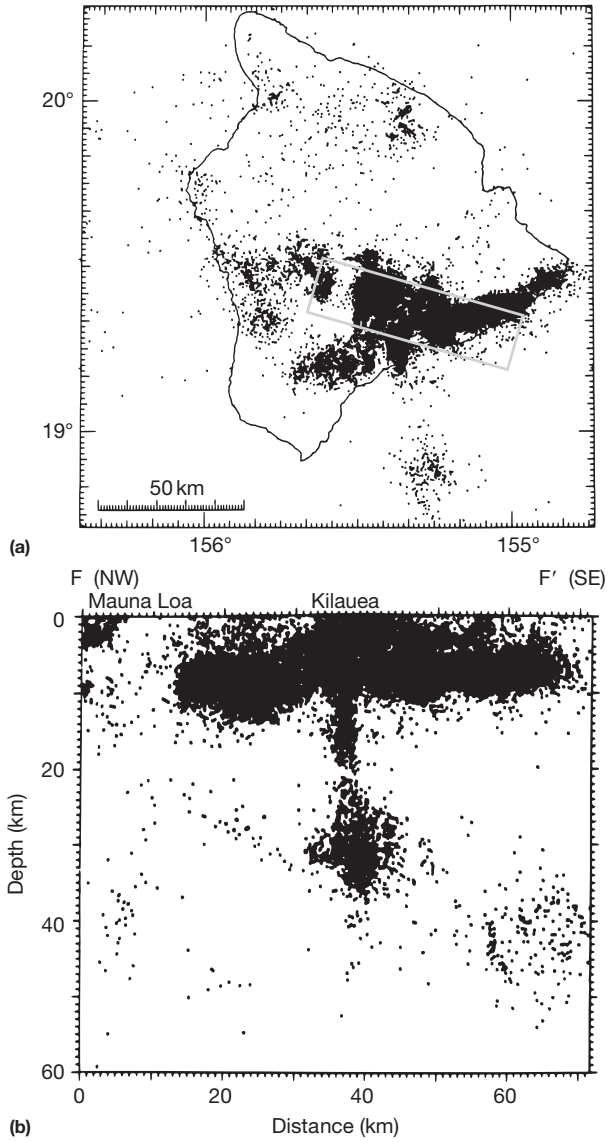


Figure 22 Seismicity of Hawaii. (a) Hypocenters of shallow earthquakes during the period 1970–84. (b) Cross section through Kilauea and Mauna Loa for earthquakes during 1970–83. The area of the cross section is that delimited by the box in the top panel. Adapted from Klein FW and Koyanagi RY (1989) The seismicity and tectonics of Hawaii. In: *The Geology of North America*, *Vol. N*: *The Eastern Pacific Ocean and Hawaii*, pp. 238–252. Boulder, CO: Geological Society of America.

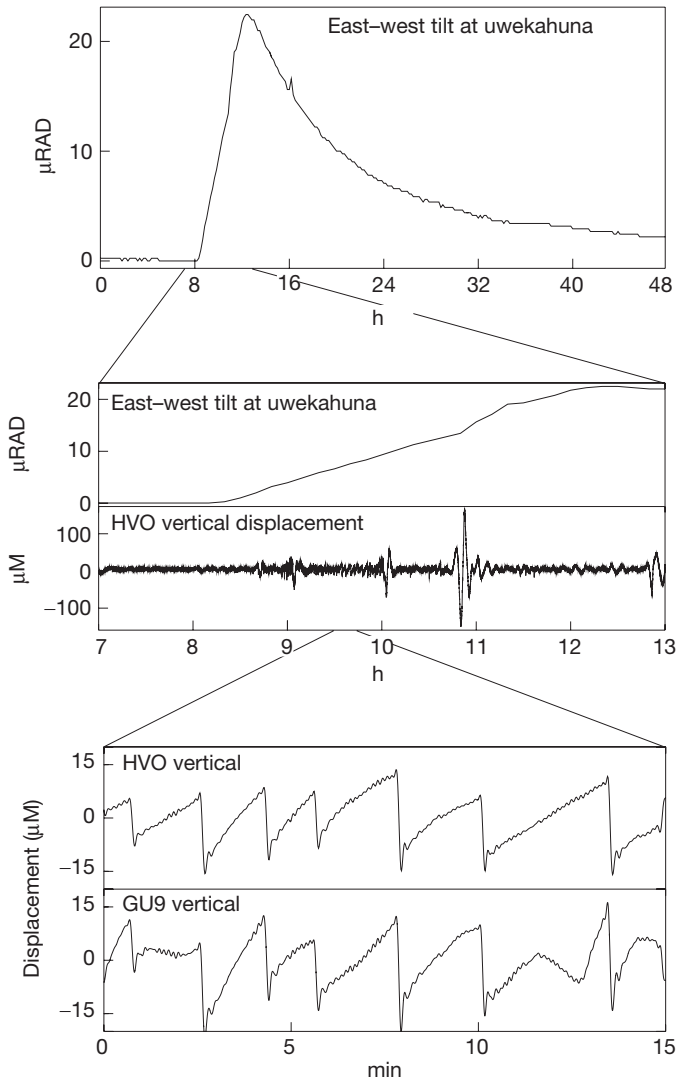


Figure 23 Tilt and broadband record observed around the summit caldera of Kilauea. Adapted from Ohminato T, Chouet B, Dawson P, and Kedar S (1998) Waveform inversion of very long period impulsive signals associated with magmatic injection beneath Kilauea Volcano, Hawaii. *Journal of Geophysical Research* 103: 23839–23862 (after Yamamoto, 2005).

4.15.6.3 Miyake

Miyake-jima is a basaltic stratovolcanic island in the Izu-Bonin arc, Japan, and has erupted quasiperiodically with an approximate interval of 20 years; the recent activities took place in 1940, 1962, and 1983. The eruption in 2000 was expected in this sense; however, what actually happened was totally unexpected: A new caldera was formed for the first time in the past 2500 years (**Figure 24**). This caldera-forming process was recorded with contemporary geophysical equipments.

Figure 25 shows the seismicity observed during the 2000 Miyake-jima activity (Sakai et al., 2001). The earthquake swarm started right beneath the volcano and migrated to the southwest (Uhira et al. (2005), not well-resolved in **Figure 25**), which may correspond to the subsurface migration of magma beneath the volcano. The swarm activity next shifted to northwest of the island, indicating a major subsurface magma migration from the volcano to the outside dike system. After a small summit eruption, the caldera formation took place, and a number of LF earthquakes are observed beneath the summit area. It appears that sucking of magma from the magma storage system beneath the volcano to the northwest dike system and the resulting magma vacancy in the volcano are the direct causes of the caldera formation (Furuya et al., 2003).

During roughly 40 days of the caldera formation, a number of peculiar bell-shaped LP (50 s) seismic (velocity) signals, as well as steps in tiltmeters, are observed once or twice a day (**Figure 26**). Based on the moment tensor inversion of the near-field and far-field broadband seismic records, Kumagai et al. (2001) associated these LP velocity pulses with inflations of a nearly vertical crack-like magma chamber due to its pressure increase as response to the repetitive stick-slip-type falling of a piston-like mass within the conduit that results in the caldera formation at the summit. Analyzing the signals of the same sequence of events recorded by the tiltmeters and

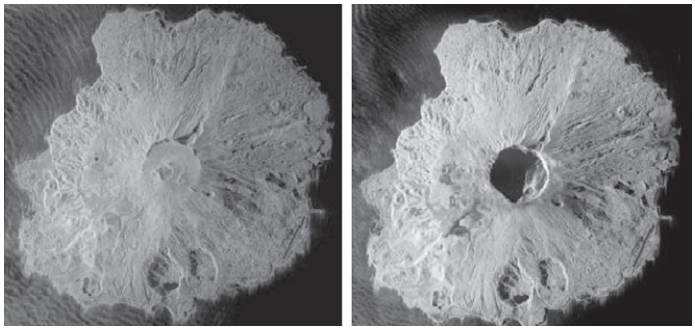


Figure 24 Airborne SAR images of Miyake-jima Volcano before and after the 2000 caldera formation. Adapted from Kumagai H, Ohminato T, Nakano M, et al. (2001) Very-long-period seismic signals and caldera formation at Miyake Island, Japan. *Science* 293: 687–690.

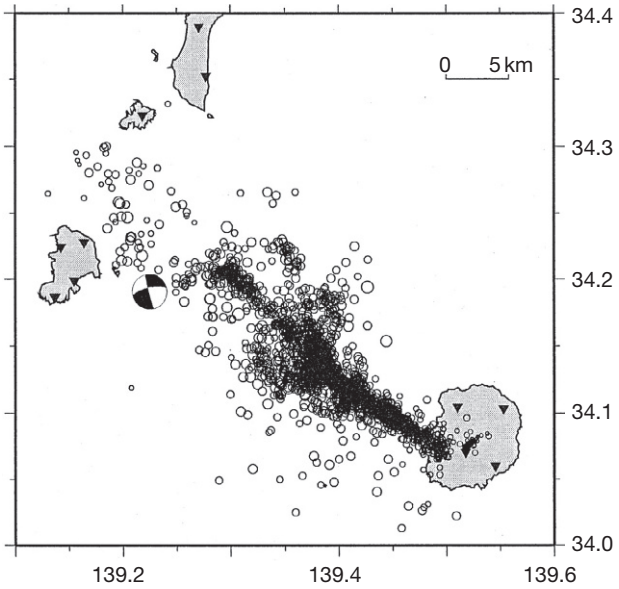


Figure 25 The seismicity of the 2000 Miyake-jima activity. Adapted from Sakai S, Yamada T, Ide S, et al. (2001) Magma migration from the point of view of seismic activity in the volcanism of Miyake-jima island in 2000. *Journal of Geography* 110: 145–155.

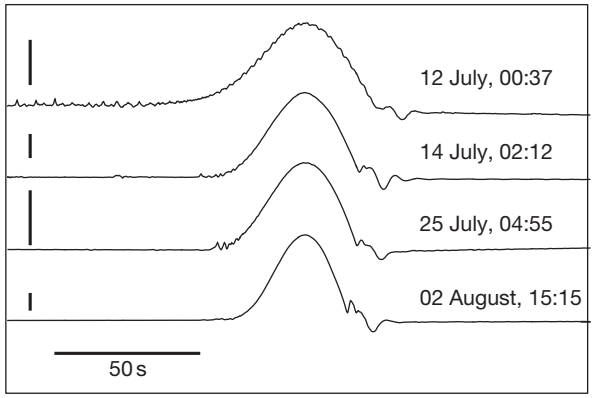


Figure 26 Vertical-component near-field velocity waveforms observed in the Miyake-jima island. Adapted from Kumagai H, Ohminato T, Nakano M, et al. (2001) Very-long-period seismic signals and caldera formation at Miyake Island, Japan. *Science* 293: 687–690.

broadband seismometers deployed in the island, Fujita et al. (2004) arrived at a different conclusion. They suggested that these events are caused by a cyclic expansion of a subsurface sill-like magma plumbing system and introduced a model based on a two-phase flow instability, called a pressure drop oscillation as the mechanism for the repeated activity. Either proposed model seems to explain a part of the observations, but not all. Although the available data may not be able to eventually resolve the difference, nevertheless, the 2000 Miyake-jima activity has provided us a rare opportunity to observe a caldera-forming event with contemporary geophysical instruments.

4.15.6.4 Stromboli

Stromboli is one of the Aeolian islands of Italy and has been almost continuously in eruption for more than at least 2000

years. In contrast to the widely used term 'Strombolian eruption' that is undiscerningly used to describe a variety of volcanic eruptions, the activity of Stromboli is characterized by short-term explosive bursts of lava ejected into the air. Since they eject relatively viscous basaltic lava, building up of the gas pressure is required to fragment the magma and results in intermittent episodic explosions.

From the seismic point of view, such activities generate a considerable number of explosion quakes, persistent volcanic tremors, and LP signals. One of the most characteristic features of seismic activities at Stromboli is the extremely wide frequency contents of these signals (**Figure 27**). Spindle-shaped signals superimposed on background continuous tremor are generated by summit eruptions, and the nature of these short-period signals has been extensively studied by many researchers (e.g., Del Pezzo, 1992; Del Pezzo et al., 1974; Ripepe and Gordeev, 1999; Saccorotti and Pezzo, 2000). These studies shed new light on the seismic activity of the volcano, which is mainly characterized by a very shallow seismicity. Such very shallow origin of the seismic signals is also confirmed by the analyses of seismic array observations (e.g., Chouet et al., 1997; Saccorotti and Pezzo, 2000), and the sources of explosion quakes and tremors are located at depths shallower than 200 m beneath the summit crater.

Recent observations using broadband seismometers reveal the existence of VLP signals with a period of about 10 s, which are accompanied with explosions, and give additional constraints on the understanding of Strombolian activities (e.g., Chouet et al., 2003, 2008; Neuberg et al., 1994). Analyses of broadband seismograms using the independent component analysis method indicate that further longer-period signals (30–40 s) are generated at almost the same region as the VLP signals (De Martino et al., 2005).

Based on various field observations and laboratory experiments, Ripepe and Gordeev (1999) and Ripepe et al. (2001) arrived at a model of Strombolian explosions that is reminiscent of the analog experimental model by Jaupart and Vergniolle (1989), in which dissolved gas bubbles are trapped at the roof of a magma reservoir as a foam layer and the periodical collapse of the layer causes ascent of gas slugs through a vertical conduit connecting the reservoir and the surface vent. The growth, flow, and burst of the coalescent gas bubbles observed in their laboratory experiments well capture the characteristics of sequential occurrence of VLP/short-period seismic signals, acoustic signals, and thermal/light emissions observed at Stromboli. **Figure 28** shows such a direct link between gas flux, magma volume flux, and seismicity, where the balance between gas/magma flux and gas overpressurization appears to determine the type of activities (i.e., effusive or explosive) (Ripepe et al., 2005). The common trend in the VLP event rate and the SO_2 emission rate during the effusive phase suggest that the rate of gas flux controls the frequency of foam coalescence and the development of gas slugs, which generate VLP signals; the decrease in SO_2 and VLP rates coincides with the increase in the explosion rate at the summit measured by the count of thermal transients and in the amplitude of acoustic pressure; this may be understood in terms of the reduction of gas/magma supplies that results in sealing of eruptive fractures and the upward migration of magma.

Performing detailed analyses of the broadband network data, Chouet et al. (2003) concluded that the source mechanism of VLP events associated with Strombolian explosions

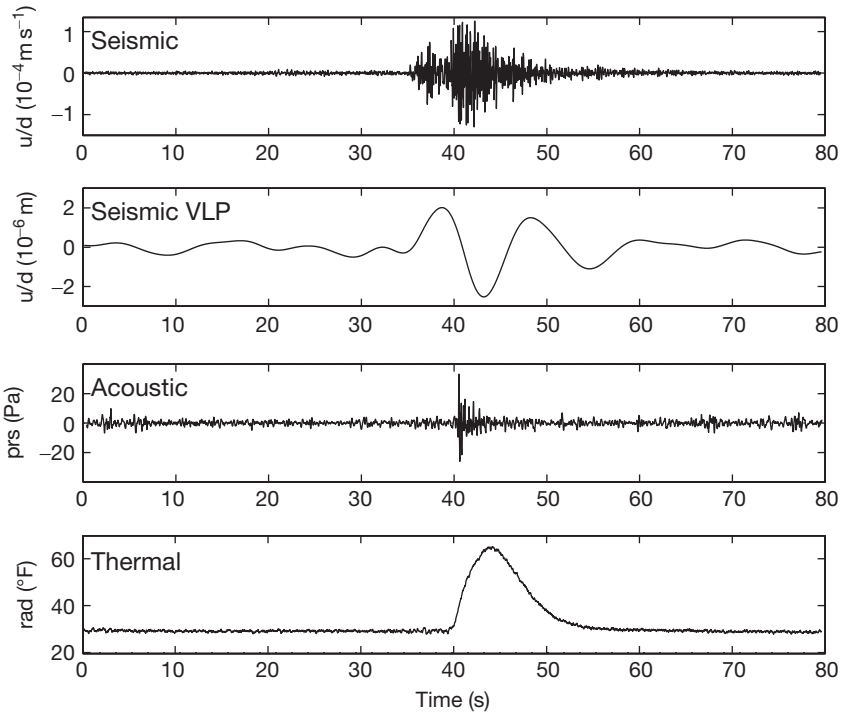


Figure 27 Example of the synchronous occurrence of seismic, acoustic, and thermal signals observed in Stromboli. From Ripepe et al., unpublished data.

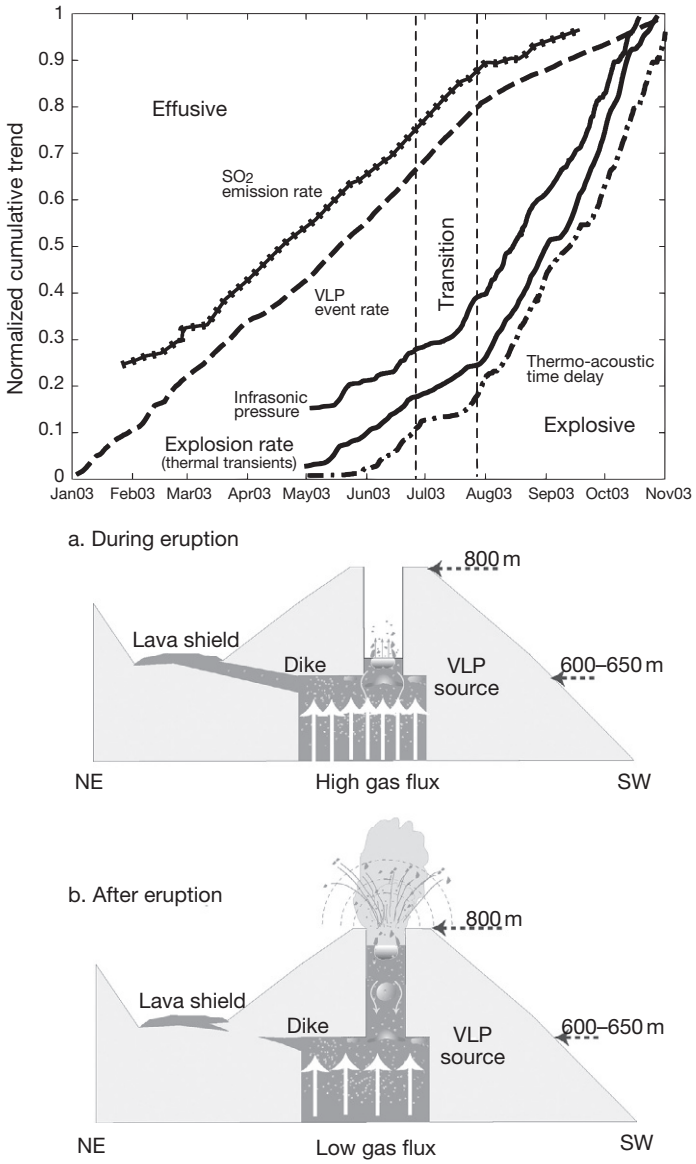


Figure 28 Example of the relation between gas flux rate and seismic/acoustic activities. During the lava effusion phase (model A), few explosions at the summit were observed while the VLP seismicity well correlated with the SO_2 emission rate. On the other hand, in the explosive phase (model B), magma rose up in the conduit due to the sealing of eruptive fractures, resulting in the gas overpressurization and the explosive activity at the summit craters. Adapted from Ripepe M, Marchetti E, Ulivieri G, et al. (2005) Effusive to explosive transition during the 2003 eruption of Stromboli volcano. *Geology* 33: 341–344.

consists of both moment tensor and single-force components, which, respectively, correspond to the opening/closing of an inclined crack just beneath the active vents and the magma flow in the conduit caused by the piston-like rise of a gas slug. To investigate the behavior of coalescence and ascent of gas slugs in such an inclined conduit, James et al. (2004) carried out laboratory experiments of the two-phase flow in vertical and inclined pipes and demonstrated that the inclination of the conduit may play an important role in controlling the flow types and size/velocity distributions of gas slugs in the fluid-filled conduit.

In spite of the difference between models in details, these results suggest that understanding of the bubble dynamics may be essential to elucidate the source process of the seismic signals and the dynamics of Strombolian activities.

4.15.7 Other Volcano-Specific Issues

4.15.7.1 Explosion Quakes

Explosion quakes are observed when explosive eruptions occur (e.g., Figures 3 and 4). An explosive volcanic eruption (Vulcanian type) can be macroscopically understood as the deflation process of a reservoir system beneath a volcano, which may be observed by geodetic (e.g., Ishihara, 1990) or broadband seismic (e.g., Figure 6) means. The source

mechanism of the accompanied explosion quake may be represented by a combination of a single force and an implosive moment tensor (**Figure 11**). Since the pioneering work of Kanamori et al. (1984), there are now a number of observations to quantify the force mechanism and size of volcanic explosions (e.g., Chouet et al., 2005; Nishimura et al., 1995; Ohminato et al., 2006).

Explosion quakes are extensively studied at Sakurajima Volcano, Japan; comparing seismic and video records, Ishihara (1985) demonstrated that explosion quakes occur at a depth of 1–2 km beneath the active crater about a few seconds prior to the explosive eruptions at the summit, which generate air shock waves. Uhira and Takeo (1994) showed that an explosive and implosive cylindrical moment tensor component at the source depth is the dominating source mechanism for the explosion quakes. By carefully analyzing borehole seismic network records, Tameguri et al. (2002) were able to decompose explosion quakes into two processes: The first one is an isotropic expansion followed by a contraction of a cylinder at a depth of 2 km, and the second is an isotropic expansion and subsequent horizontal contraction at depths of 0.25–0.5 km beneath the crater bottom, which occur about 1 s after the onset of the first and coincide with the generation of the shock wave. The plausibility of this second process was numerically demonstrated by Nishimura and Chouet (2003), who considered the behavior of a reservoir–conduit system embedded in a homogeneous crust, which is filled with a compressive fluid; it was shown that an abrupt pressure increase occurs just beneath the vent, which can generate pulse-like Rayleigh waves as observed by Tameguri et al. (2002), and that the strength of the lid played an important role in defining eruption types.

4.15.7.1.1 Scaling law of volcanic explosion

Nishimura (1998) showed that the far-field displacement due to an explosion is proportional to the area (square of radius) of the crater vent and obtained a scaling law of volcanic explosion quakes, as similar to that of earthquakes. **Figure 29** compares the vent radii and estimated seismic magnitudes of the largest event for studied volcanoes by various investigators; it shows that the magnitude is essentially proportional to the cross-sectional area of the vent for vent radii ranging from 10 to 600 m, as predicted by the theory, and that the initial excess pressure in the reservoir is of the order of a few megapascals, a similar value as the stress drops of earthquakes. This suggests that the size of explosion quakes may be estimated by knowing the vent radius, although it does not necessarily imply that it is possible to predict the size of future eruptions, as it is not guaranteed that the size of present vent is the same for the future ones.

4.15.7.2 Triggered Seismicity

It has been well known among local observers that volcanoes are 'sensitive,' as seismicity there often shows changes apparently associated with external disturbances (tides or large earthquakes). For example, Sassa (1936) documented diurnal and semidiurnal periodicities in the occurrence of volcanic seismicity of Aso Volcano. McNutt and Beavan (1981) also suggested a tidal correlation of seismicity at Pavlof Volcano (cf., Emter, 1997).

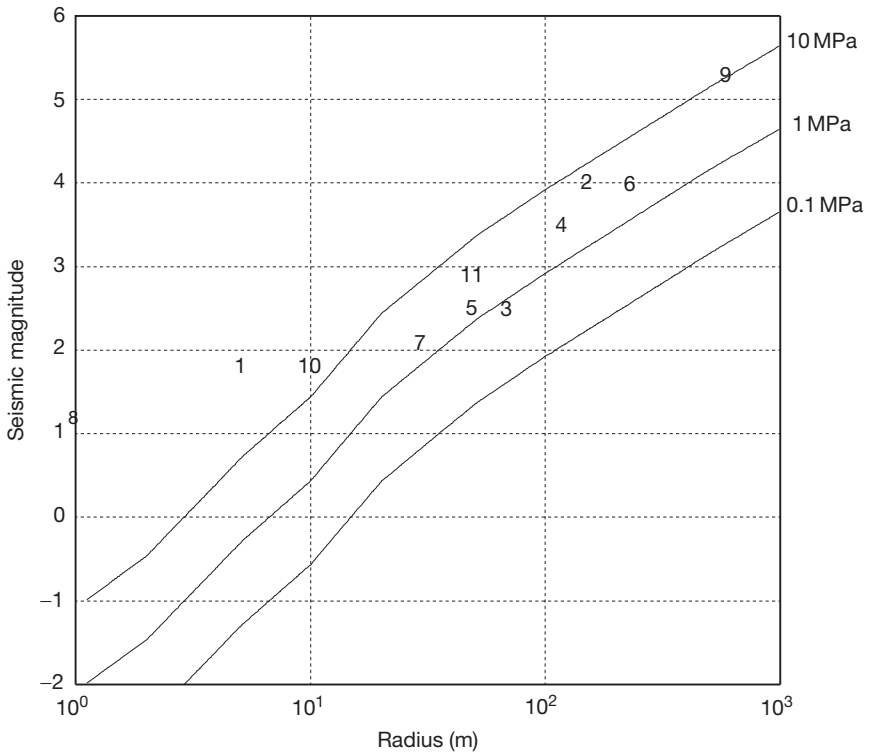


Figure 29 Relation between the seismic magnitude of eruption quakes and the observed vent radius. Numbers from 1 to 11 indicate representative points of different volcanoes. Lines show theoretical relation for assumed initial excess pressures. Adapted from Nishimura T (1998) Source mechanisms of volcanic explosion earthquakes: Single force and implosive sources. *Journal of Volcanology and Geothermal Research* 86: 97–106.

There are also convincing lines of evidence that the passage of large amplitude seismic waves dynamically triggers seismicity, especially at volcanoes. Hill et al. (1993) documented the seismicity increase in Long Valley Caldera, USA, following the Mw = 7.3 1992 Landers earthquake, California (also Gomberg et al., 2001; Prejean et al., 2004). Brodsky and Prejean (2005) made a systematic survey of remotely triggered seismicity at Long Valley; showing the evidence for a frequency-dependent triggering threshold, they concluded that LP (>30 s) waves are more effective at generating local seismicity than short-period waves of comparable magnitude; among the several proposed mechanisms for such remote dynamic triggering, they suggested that the mechanism that invokes fluid flow through a porous medium can produce such a frequency dependence of triggering threshold (~5 kPa). This mechanism also seems to explain why volcanoes are sensitive to external disturbances, as unclogging fractures to drive fluid flow (Brodsky et al., 2003) may be easily initiated in volcanic/geothermal areas where the constant precipitation from volcanic fluids generate fragile blockages in fractures and faults. West et al. (2005) observed periodically (20–30 s) triggered seismicity at Mt. Wrangell, Alaska, when LP Rayleigh waves generated by the Mw = 9.0 2004 Sumatra earthquake swept across Alaska, and concluded that the extensional stresses of ~25 kPa due to the Rayleigh waves triggered shear failure on normal faults, which may be also influenced by the aforementioned fluid pumping mechanism. These observations suggest that the presence of ample fluid in volcanic systems makes them ideal experimental field for elucidating the earthquake-triggering mechanism.

4.15.7.3 Seismic Wave Scattering in Volcanoes

Dynamic interactions between volcanic fluids and volcanic edifice along with more large-scale tectonic activities make the volcanic structure very complex, and volcano is considered to be one of the most heterogeneous parts of the Earth's crust. The understanding of inhomogeneities in volcanoes may thus provide us important information on the various volcanic processes as well as the volcanic structure. One way to infer the distribution and movement of volcanic fluids and resultant heterogeneities beneath volcanoes is the use of tomography for seismic velocity and attenuation, which basically relies on the coherent waves (mainly direct waves). The tomographic method, however, has some inherent limitations in its resolution partly due to the wavelength of the waves used in the analyses, and it is rather difficult to image small-scale heterogeneities like intruded dykes and fractures using the method.

An alternative approach to image small-scale heterogeneities is the use of scattered waves. From this point of view, the scattering properties and their spatial variation around active volcanoes have been studied based on various kinds of wave scattering model, and it has been revealed that small-scale heterogeneity beneath volcanoes is one or two orders stronger than that of ordinal crust (e.g., Sato et al., 2012). Such strong heterogeneity causes the multiple scattering of seismic waves and mode conversion between P- and S-waves, and thus, the waveforms of volcanic earthquakes often exhibit spindle-like envelopes having small direct waves and long-lasting coda waves. While these complex waveforms, especially those missing clear coherent phases, seem not suitable for conventional analyses, recent progresses in theoretical and observational researches on seismic interferometry (e.g., Schuster et al., 2004) reveal that multiply scattered waves instead have a significant advantage that the slight change in the elastic properties of medium, which have little influence on the first arrivals, are amplified and thus readily observed. In addition, the existence of dominant mode conversion due to strong heterogeneity in volcanic environment (e.g., Yamamoto and Sato, 2010) makes the situation more ideal for the application of the seismic interferometry. The method has been effectively applied to monitor the slight medium changes due to magmatic activities at active volcanoes (e.g., Brenguier et al., 2008; Nagaoka et al., 2010), and it may be one of the most promising techniques to monitor the movement of volcanic fluids in volcanic systems and predict volcanic eruptions. Future researches on the elementary physics of seismic wave scattering including the resonant scattering of small fluid inclusions (e.g., Margerin, 2013) may also contribute to better understanding of behavior of volcanic fluids.

4.15.8 Concluding Remarks

One of the ultimate goals of volcano seismology is to extract quantitative information about the volcanic fluid transport systems beneath active volcanoes and to utilize such information for forecasting short- and long-term eruptive behaviors. As described in this chapter, macroscopic behaviors of dynamic interactions between volcanic fluids and volcanic edifices have been extensively studied in volcano seismology, and significant progress has been made; the observational/numerical/experimental studies have established firm foundations for explaining the macroscopic processes responsible for generating observed volcanic seismic signals and other related volcanic activities. Considering the rapid progress in technologies for the field observations and numerical simulations, the future of these lines of studies seems to be promising; the accumulation of high-quality digital data at different volcanoes around the world will further bring significant advances in volcano seismology, as well as volcanology in general; the increase in computer power remarkably extends our capability to simulate complicated dynamic interactions between phases and enables us to identify and separate distinct signatures of volcanic seismic signals; the recent advance in high-sampling, high-precision GPS and InSAR observations may also help our understanding from the observational aspect.

Nevertheless, our understanding of the elementary physics of multiphase systems, such as the dynamics of bubbles in magmatic fluids and their flows, is still immature, although such understanding fused with the established studies is indispensable for the elucidation of the whole volcanic fluid systems. For instance, recently, James et al. (2006) experimentally demonstrated that the geometry of the conduit is the key parameter governing the flow pattern of ascending slugs in viscous magma, which act as a trigger/source of volcanic seismic signals. Such a parameter of the conduit system may be retrieved by acoustic measurements as demonstrated by Garcés et al. (2000), and the integration of these studies will provide a great feedback to volcano seismology. Time-varying

moment-tensor and single-force components estimated from high-quality seismic data may also be an important clue to image nonlinear fluid dynamics in volcanic conduits. Such studies on the modeling of physical processes, in cooperation with multiparameter observations and numerical simulations, may bring a new boost to volcano seismology in the near future. Volcano seismology, thus, remains to be one of the most challenging fields in Earth science as an integrated science of multidisciplinary researches on multiscale/multiphase systems.

Acknowledgments

The authors thank Hiroo Kanamori for his patience and encouragement as the editor eventually to complete this manuscript and Takao Ohminato, Hiroyuki Kumagai, and Bernard Chouet for the valuable comments on the original manuscript for the first edition. As for the second edition, we made several minor modifications in the text with updated references. Also, **Section 4.15.7.3 'Seismic Wave Scattering in Volcanoes'** is newly added.

References

Aki K (1957) Space and time spectra of stationary stochastic waves with special reference to microtremors. *Bulletin of the Earthquake Research Institute, University of Tokyo* 35: 415–456.

Aki K (1984) Evidence for magma intrusion during the mammoth lakes earthquakes of May 1980 and implications of the absence of volcanic (harmonic) tremor. *Journal of Geophysical Research* 89: 7689–7696.

Aki K, Fehler M, and Das S (1977) Source mechanism of volcanic tremor: Fluid-driven crack models and their application to the 1963 Kilauea eruption. *Journal of Volcanology and Geothermal Research* 2: 259–287.

Aki K and Koyanagi R (1981) Deep volcanic tremor and magma ascent mechanism under Kilauea, Hawaii. *Journal of Geophysical Research* 86: 7095–7109.

Aki K and Lee WHK (2003) Appendix 1 – Glossary of interest to earthquake and engineering seismologists. In: *International Handbook of Earthquake and Engineering Seismology, Part B*, pp. 1793–1856. New York: Academic Press.

Aki K and Richards PG (2002) *Quantitative Seismology*, 2nd edn. Sausalito, CA: University Science Books.

Alidibirov M and Dingwell DB (1996) Magma fragmentation by rapid decompression. *Nature* 380: 146–149.

Almendros J, Chouet B, and Dawson P (2001) Spatial extent of a hydrothermal system at Kilauea Volcano, Hawaii, determined from array analyses of shallow long-period seismicity: 2. Results. *Journal of Geophysical Research* 106: 13581–13598.

Almendros J, Chouet B, Dawson P, and Bond T (2002) Identifying elements of the plumbing system beneath Kilauea Volcano, Hawaii, from the source locations of very-long-period signals. *Geophysical Journal International* 148: 303–312.

Almendros J, Chouet B, Dawson P, and Huber C (2002) Mapping the sources of the seismic wavefield at Kilauea Volcano, Hawaii, using data recorded on multiple seismic antennas. *Bulletin of Seismological Society of America* 92: 2333–2351.

Aoyama H and Oshima H (2007) Tilt change recorded by broadband seismometer prior to small phreatic explosion of Meakan-dake volcano, Hokkaido, Japan. *Geophysical Research Letters* 35: L06307.

Aoyama H and Takeo M (2001) Wave properties and focal mechanisms of N-type earthquakes at Asama volcano. *Journal of Volcanology and Geothermal Research* 105: 163–182.

Arciniega-Ceballos A, Chouet B, and Dawson P (1999) Very long-period signals associated with Vulcanian explosions at Popocatépetl Volcano, Mexico. *Geophysical Research Letters* 26: 3013–3016.

Arciniega-Ceballos A, Chouet B, and Dawson P (2003) Long-period events and tremor at Popocatepetl volcano (1994–2000) and their broadband characteristics. *Bulletin of Volcanology* 65: 124–135.

Arciniega-Ceballos A, Chouet B, Dawson P, and Asch G (2008) Broadband seismic measurements of degassing activity associated with lava effusion at Popocatépetl Volcano, Mexico. *Journal of Volcanology and Geothermal Research* 170: 12–23.

Aster R, Mah S, Kyle P, et al. (2003) Very long period oscillations of Mount Erebus Volcano. *Journal of Geophysical Research* 108: 2522.

Aster R, Zandomeneghi D, Mah S, et al. (2008) Moment tensor inversion of very long period seismic signals from Strombolian eruptions of Erebus Volcano. *Journal of Volcanology and Geothermal Research* 177: 635–647.

Auger E, D'Auria L, Martini M, Chouet B, and Dawson P (2006) Real-time monitoring and massive inversion of source parameters of very long period seismic signals: An application to Stromboli volcano. *Geophysical Research Letters* 33: L04301.

Backus G and Mulcahy M (1976) Moment tensors and other phenomenological descriptions of seismic sources, I. Continuous displacements. *Geophysical Journal of the Royal Astronomical Society* 46: 341–361.

Balmforth EJ, Craster RV, and Rust AC (2005) Instability in flow through elastic conduits and volcanic tremor. *Journal of Fluid Mechanics* 527: 353–377.

Ben-Menahem A and Singh SJ (1981) *Seismic Waves and Sources*. Berlin: Springer-Verlag.

Benoit JP and McNutt SR (1997) New constraints on source processes of volcanic tremor at Arenal Volcano, Costa Rica, using broadband seismic data. *Geophysical Research Letters* 24: 449–452.

Biot MA (1952) Propagation of elastic waves in a cylindrical bore containing a fluid. *Journal of Applied Physics* 23: 997–1005.

Biot MA (1956) The theory of propagation of elastic waves in a fluid-saturated porous solid. I. Low-frequency range. *Journal of the Acoustical Society of America* 28: 168–178.

Brenguier F, Shapiro NM, Campillo M, et al. (2008) Towards forecasting volcanic eruptions using seismic noise. *Nature Geoscience* 1: 126–130.

Brodsky EE and Prejean SG (2005) New constraints on mechanisms of remotely triggered seismicity at Long Valley Caldera. *Journal of Geophysical Research* 110: B04302.

Brodsky EE, Roeloffs E, Woodcock D, Gall I, and Manga M (2003) A mechanism for sustained groundwater pressure changes induced by distant earthquakes. *Journal of Geophysical Research* 108: 2390.

Brodsky EE, Sturtevant B, and Kanamori H (1998) Earthquakes, volcanoes, and rectified diffusion. *Journal of Geophysical Research* 103: 23827–23838.

Chouet B (1985) Excitation of a buried magmatic pipe: A seismic source model for volcanic tremor. *Journal of Geophysical Research* 90: 1881–1893.

Chouet B (1986) Dynamics of a fluid-driven crack in three dimensions by the finite difference method. *Journal of Geophysical Research* 91: 13967–13992.

Chouet B (1992) A seismic model for the source of long-period events and harmonic tremor. In: *Volcanic Seismology*, pp. 133–156. Berlin: Springer-Verlag.

Chouet B (1996a) Long-period volcano seismicity: Its source and use in eruption forecasting. *Nature* 380: 309–316.

Chouet B (1996b) New methods and future trends in seismological volcano monitoring. In: *Monitoring and Mitigation of Volcano Hazards*, pp. 23–97. New York: Springer-Verlag.

Chouet B (2003) Volcano seismology. *Pure and Applied Geophysics* 160: 739–788.

Chouet B and Dawson P (2011) Shallow conduit system at Kilauea Volcano, Hawaii, revealed by seismic signals associated with degassing bursts. *Journal of Geophysical Research* 116: B12317.

Chouet B, Dawson P, and Arciniega-Ceballos A (2005) Source mechanism of Vulcanian degassing at Popocatepetl Volcano, Mexico, determined from waveform inversions of very long period signals. *Journal of Geophysical Research* 110: B07301.

Chouet BA, Dawson PB, James MR, and Lane SJ (2010) Seismic source mechanism of degassing bursts at Kilauea Volcano, Hawaii: Results from waveform inversion in the 10–50 s band. *Journal of Geophysical Research* 115: B09311.

Chouet B, Dawson P, and Martini P (2008) Shallow-conduit dynamics at Stromboli Volcano, Italy, imaged from waveform inversion. In: *Fluid Motions in Volcanic Conduits: A Source of Seismic and Acoustic Signals*, pp. 57–84. London: Geological Society.

Chouet B, Dawson P, Ohminato T, et al. (2003) Source mechanisms of explosions at Stromboli Volcano, Italy, determined from moment-tensor inversions of very-long-period data. *Journal of Geophysical Research* 108: 2019.

Chouet B and Julian BR (1985) Dynamics of an expanding fluid-filled crack. *Journal of Geophysical Research* 90: 11187–11198.

Chouet B, Luca GD, Milana G, Dawson PB, Martini M, and Scarpa R (1998) Shallow velocity structure of Stromboli Volcano, Italy, derived from small-aperture array measurements of Strombolian tremor. *Bulletin of Seismological Society of America* 88: 653–666.

Chouet B, Page RA, Stephens CD, Lahr JC, and Power JA (1994) Precursory swarms of long-period events at Redoubt Volcano (1989–1990), Alaska: Their origin and use as a forecasting tool. *Journal of Volcanology and Geothermal Research* 62: 95–135.

Chouet B, Saccorotti G, Martini M, et al. (1997) Source and path effects in the wave fields of tremor and explosions at Stromboli Volcano, Italy. *Journal of Geophysical Research* 102: 15129–15150.

Corona-Romero P, Arciniega-Ceballos A, and Snchez-Sesma F (2012) Simulation of LP seismic signals modeling the fluid-rock dynamic interaction. *Journal of Volcanology and Geothermal Research* 211–212: 92–111.

Crosson RS and Bame DA (1985) A spherical source model for low frequency volcanic earthquake. *Journal of Geophysical Research* 90: 10237–10247.

Dahm T (1996) Relative moment tensor inversion based on ray theory: Theory and synthetic tests. *Geophysical Journal International* 124: 245–257.

Dawson P, Dietel C, Chouet B, Honma K, Ohminato T, and Okubo P (1998) A digitally telemetered broadband seismic network at Kilauea Volcano, Hawaii. U.S. Geological Survey, Open File Report 98–108, 1–121.

Dawson P, Whilldin D, and Chouet B (2004) Application of near real-time radial semblance to locate the shallow magmatic conduit at Kilauea Volcano, Hawaii. *Geophysical Research Letters* 31: L21606.

De Luca G, Scarpa R, Pezzo ED, and Simini M (1997) Shallow structure of Mt. Vesuvius volcano, Italy, from seismic array analysis. *Geophysical Research Letters* 24: 481–484.

De Martino S, Falanga M, Scarpa R, and Godano C (2005) Very-long-period volcanic tremor at Stromboli, Italy. *Bulletin of the Seismological Society of America* 95: 1186–1192.

Del Pezzo E (1992) Wave polarization and location of the source of the explosion quakes at Stromboli volcano. In: *Volcanic Seismology*, pp. 279–296. Heidelberg: Springer-Verlag.

Del Pezzo E, Guerra I, Bascio AL, Luongo G, Nappi G, and Scarpa R (1974) Microtremors and volcanic explosions at Stromboli – Part 2. *Bulletin of Volcanology* 38: 1023–1036.

Dreger DS, Tkali H, and Johnston M (2000) Dilational processes accompanying earthquakes in the Long Valley Caldera. *Science* 288: 122–125.

Dreier R, Widmer R, Schick R, and Zürn W (1994) Stacking of broad-band seismograms of shocks at Stromboli. *Acta Vulcanologica* 5: 165–172.

Dunham EM and Ogden DE (2012) Guided waves along fluid-filled cracks in elastic solids and instability at high flow rates. *Journal of Applied Mechanics* 79: 031020.

Dziewonski AM, Chou TA, and Woodhouse JH (1981) Determination of earthquake source parameters from waveform data for studies of global and regional seismicity. *Journal of Geophysical Research* 86: 2825–2852.

Eissler HK and Kanamori H (1987) A single-force model for the 1975 Kalapana, Hawaii, earthquake. *Journal of Geophysical Research* 92: 4827–4836.

Ekström G, Nettles M, and Abers GA (2003) Glacial earthquakes. *Science* 302: 622–624.

Emter D (1997) Tidal triggering of earthquakes and volcanic events. In: *Tidal Phenomena, Lecture Notes in Earth Sciences*, pp. 293–309. Berlin: Springer.

Eshelby JD (1957) The determination of the elastic field of an ellipsoidal inclusion, and related problems. *Proceedings. Royal Society of London* A241: 376–396.

Falsaperla S, Langer H, Martinelli B, and Schick R (1994) Seismic measurements on Stromboli volcano in a wide frequency range. *Acta Vulcanologica* 5: 173–178.

Ferrazzini V and Aki K (1987) Slow waves trapped in a fluid-filled infinite crack: Implication for volcanic tremor. *Journal of Geophysical Research* 92: 9215–9223.

Ferrazzini V, Aki K, and Chouet B (1991) Characteristics of seismic waves composing Hawaiian volcanic tremor and gas-piston events observed by a near-source array. *Journal of Geophysical Research* 96: 6199–6209.

Ferrick M, Qamar A, and Lawrence WFS (1982) Source mechanism of volcanic tremor. *Journal of Geophysical Research* 87: 8675–8683.

Foulger GR, Julian BR, Hill DP, Pitt AM, Malin PE, and Shalev E (2004) Non-double-couple microearthquakes at Long Valley caldera, California, provide evidence for hydraulic fracturing. *Journal of Volcanology and Geothermal Research* 132: 45–71.

Fujita E (2008) Banded tremor at Miyakejima volcano, Japan: Implication for two-phase flow instability. *Journal of Geophysical Research* 113: B04207.

Fujita E and Ida Y (2003) Geometrical effects and low-attenuation resonance of volcanic fluid inclusions for the source mechanism of long-period earthquakes. *Journal of Geophysical Research* 108: 2118.

Fujita E, Ida Y, and Oikawa J (1995) Eigen oscillation of a fluid sphere and source mechanism of harmonic volcanic tremor. *Journal of Volcanology and Geothermal Research* 69: 365–378.

Fujita E, Ukawa M, and Yamamoto E (2004) Subsurface cyclic magma sill expansions in the 2000 Miyakejima. *Journal of Geophysics* 109: B04205.

Furuya M, Okubo S, Sun W, et al. (2003) Spatiotemporal gravity changes at Miyakejima Volcano, Japan: Caldera collapse, explosive eruptions and magma movement. *Journal of Geophysical Research* 108: 2219.

Garcés M, McNutt SR, Hansen RA, and Eichelberger JC (2000) Application of wave-theoretical seismoacoustic models to the interpretation of explosion and eruption tremor signals radiated by Pavlof volcano, Alaska. *Journal of Geophysical Research* 105: 3039–3058.

Gil Cruz F and Chouet B (1997) Long-period events, the most characteristic seismicity accompanying the emplacement and extrusion of a lava dome in Galeras Volcano, Colombia, in 1991. *Journal of Volcanology and Geothermal Research* 77: 121–158.

Gilbert F (1970) Excitation of the normal-modes of the Earth by earthquake sources. *Geophysical Journal of the Royal Astronomical Society* 23: 119–123.

Goldstein P and Archuleta RJ (1991) Deterministic frequency-wavenumber methods and direct measurements of rupture propagation during earthquakes using a dense array: Theory and methods. *Journal of Geophysical Research* 96: 6173–6185.

Goldstein P and Chouet B (1994) Array measurements and modeling of sources of shallow volcanic tremor at Kilauea Volcano, Hawaii. *Journal of Geophysical Research* 99: 2637–2652.

Gomberg J, Reasenberg PA, Bodin P, and Harris RA (2001) Earthquake triggering by seismic waves following the Landers and Hector Mine earthquakes. *Nature* 411: 462–466.

Gordeev EI, Saltykov VA, Sinitsin VI, and Chebrov VN (1990) Temporal and spatial characteristics of volcanic tremor wave fields. *Journal of Volcanology and Geothermal Research* 40: 89–101.

Graizer V (2006) Tilts in strong ground motion. *Bulletin of Seismological Society of America* 96: 2090–2102.

Hagerty MT, Schwartz SY, Garcés MA, and Protti M (2000) Analysis of seismic and acoustic observations at Arenal Volcano, Costa Rica, 1995–1997. *Journal of Volcanology and Geothermal Research* 101: 27–65.

Hasegawa HS and Kanamori H (1987) Source mechanism of the magnitude 7.2 Grand Banks earthquake of November 1929; double couple or submarine landslide? *Bulletin of Seismological Society of America* 77: 1984–2004.

Hasegawa A and Yamamoto A (1994) Deep, low-frequency microearthquakes in or around seismic low-velocity zones beneath active volcanoes in northeastern Japan. *Tectonophysics* 233: 233–252.

Hayashi Y and Morita Y (2003) An image of a magma intrusion process inferred from precise hypocentral migrations of the earthquake swarm east of the Izu Peninsula. *Geophysical Journal International* 153: 159–174.

Hellweg P, Seidl D, Brotopuspito KS, and Brustle W (1994) Team investigates activity at Mt. Semeru, Java, volcano. *EOS, Transactions American Geophysical Union* 75: 313–317.

Hidayat D, Chouet B, Voight B, Dawson P, and Ratdomopurbo A (2002) Source mechanism of very-long-period signals accompanying dome growth activity at Merapi volcano, Indonesia. *Geophysical Research Letters* 29: 2118.

Hidayat D, Voight B, Langston C, Ratdomopurbo A, and Ebeling C (2000) Broadband seismic experiment at Merapi Volcano, Java, Indonesia: Very-long-period pulses embedded in multiphase earthquakes. *Journal of Volcanology and Geothermal Research* 100: 215–231.

Hill DP, Dawson P, Johnston MJS, Pitt AM, Biasi G, and Smith K (2002) Very-long-period volcanic earthquakes beneath Mammoth Mountain, California. *Geophysical Research Letters* 29: 1370.

Hill DP and Prejean S (2005) Magmatic unrest beneath Mammoth Mountain, California. *Journal of Volcanology and Geothermal Research* 146: 257–283.

Hill DP, Reasenberg PA, Michael A, et al. (1993) Seismicity remotely triggered by the magnitude 7.3 Landers, California, earthquake. *Science* 260: 1617–1623.

Hudson JA, Pearce RG, and Rogers RM (1989) Source type plot for inversion of the moment tensor. *Journal of Geophysical Research* 94: 765–774.

Ibanez JM, Pezzo ED, Almendros J, et al. (2000) Seismovolcanic signals at Deception Island volcano, Antarctica: Wave field analysis and source modeling. *Journal of Geophysical Research* 105: 13905–13932.

Ichihara M and Brodsky EE (2006) A limit on the effect of rectified diffusion in volcanic systems. *Geophysical Research Letters* 33: L02316.

Ichihara M, Rittel D, and Sturtevant B (2002) Fragmentation of a porous viscoelastic material: Implications to magma fragmentation. *Journal of Geophysical Research* 107: 2229.

Ida Y (1996) Cyclic fluid effusion accompanied by pressure change: Implication for volcanic eruptions and tremor. *Geophysical Research Letters* 23: 1457–1460.

Ishihara K (1985) Dynamical analysis of volcanic explosion. *Journal of Geodynamics* 3: 327–349.

Ishihara K (1990) *Pressure Sources and Induced Ground Deformation Associated with Explosive Eruptions of an Andesitic Volcano, Sakurajima Volcano, Japan*. London: Wiley, pp. 335–356.

Iwamura K and Kaneshima S (2005) Numerical simulation of the steam-water flow instability as a mechanism of long-period ground vibrations at geothermal areas. *Geophysical Journal International* 163: 833–851.

James MR, Lane SJ, and Chouet BA (2006) Gas slug ascent through changes in conduit diameter: Laboratory insights into a volcano-seismic source process in low-viscosity magmas. *Journal of Geophysical Research* 111: B05201.

James MR, Lane SJ, Chouet B, and Gilbert JS (2004) Pressure changes associated with the ascent and bursting of gas slugs in liquid-filled vertical and inclined conduits. *Journal of Volcanology and Geothermal Research* 129: 61–82.
Jaupart C and Vergniolle S (1989) The generation and collapse of a foam layer at the roof of a basaltic magma chamber. *Journal of Fluid Mechanics* 203: 347–380.
Jousset P, Neuberg J, and Sturton S (2003) Modelling the time-dependent frequency content of low-frequency volcanic earthquakes. *Journal of Volcanology and Geothermal Research* 128: 201–223.
Julian BR (1983) Evidence for dyke intrusion earthquake mechanisms near Long Valley Caldera, California. *Nature* 303: 323–325.
Julian BR (1994) Volcanic tremor: Nonlinear excitation by fluid flow. *Journal of Geophysical Research* 99: 11859–11878.
Julian BR and Foulger GR (1996) Earthquake mechanisms from linear-programming inversion of seismic-wave amplitude ratios. *Bulletin of Seismological Society of America* 86: 972–980.
Julian BR, Miller AD, and Foulger GR (1997) Non-double-couple earthquake mechanisms at the Hengill-Grensdalur volcanic complex, southwest Iceland. *Geophysical Research Letters* 24: 743–746.
Julian BR, Miller AD, and Foulger GR (1998) Non-double-couple earthquakes 1. Theory. *Reviews of Geophysics* 36: 525–549.
Kamo K, Furuzawa T, and Akamatsu J (1977) Some natures of volcanic micro-tremors at the Sakurajima volcano. *Bulletin of the Volcanological Society of Japan* 22: 41–58.
Kanamori H, Ekström G, Dziewonski A, Barker JS, and Sipkin SA (1993) Seismic radiation by magma injection: An anomalous seismic event near Tori Shima, Japan. *Journal of Geophysical Research* 98: 6511–6522.
Kanamori H and Given JW (1982) Analysis of long-period seismic waves excited by the May 18, 1980, eruption of Mount St. Helens – A terrestrial monopole. *Journal of Geophysical Research* 87: 5422–5432.
Kanamori H, Given JW, and Lay T (1984) Analysis of seismic body waves excited by the Mount St. Helens eruption of May 18, 1980. *Journal of Geophysical Research* 89: 1856–1866.
Kanamori H and Mori J (1992) Harmonic excitation of mantle Rayleigh waves by the 1991 eruption of Mount Pinatubo Philippines. *Geophysical Research Letters* 19: 721–724.
Kaneshima S, Kawakatsu H, Matsubayashi H, et al. (1996) Mechanism of phreatic eruptions at Aso volcano inferred from near-field broadband seismic observations. *Science* 273: 643–645.
Kawakatsu H (1989) Centroid single force inversion of seismic waves generated by landslides. *Journal of Geophysical Research* 94: 12363–12374.
Kawakatsu H (1998) On the real-time monitoring of the long-period seismic wavefield. *Bulletin of the Earthquake Research Institute, University of Tokyo* 73: 267–274 (also in Methods and applications of signal processing in seismic network operations (2002), Springer, pp. 251–257).
Kawakatsu H, Kaneshima S, Matsubayashi H, et al. (2000) Aso94: Aso seismic observation with broadband instruments. *Journal of Volcanology and Geothermal Research* 101: 129–154.
Kawakatsu H and Montagner J-P (2008) Time-reversal seismic-source imaging and moment-tensor inversion. *Geophysical Journal International* 175: 686–688.
Kawakatsu H, Ohminato T, and Ito H (1994) 10s-period volcanic tremors observed over a wide area in southwestern Japan. *Geophysical Research Letters* 21: 1963–1966.
Kawakatsu H, Ohminato T, Ito H, et al. (1992) Broadband seismic observation at the Sakurajima volcano, Japan. *Geophysical Research Letters* 19: 1959–1962.
Kieffer SW (1977) Sound speed in liquid-gas mixtures: Water-air and water-steam. *Journal of Geophysical Research* 82: 2895–2904.
Kikuchi M and Kanamori H (1982) Inversion of complex body waves. *Bulletin of Seismological Society of America* 72: 491–506.
Klein FW and Koyanagi RY (1989) The seismicity and tectonics of Hawaii. *The Geology of North America*, vol. N, pp. 238–252. Boulder, CO: Geological Society of America.
Knopoff L and Randall M (1970) The compensated linear-vector dipole: A possible mechanism for deep earthquakes. *Journal of Geophysical Research* 75: 4957–4963.
Kobayashi T, Ohminato T, Ida Y, and Fujita E (2009) Very long period seismic signals observed before the caldera formation with the 2000 Miyake-jima volcanic activity, Japan. *Journal of Geophysical Research* 114: B02211.
Krauklis PV (1962) On low-frequency oscillations of a liquid layer in an elastic medium. *Prikladnaya Matematika i Mekhanika (Journal of Applied Mathematics and Mechanics)* 26: 1111–1115.
Kubotera A (1974) Volcanic tremors at Aso volcano. In: *Physical volcanology*, pp. 29–47. Amsterdam: Elsevier.
Kumagai H (2006) Temporal evolution of a magmatic dike system inferred from the complex frequencies of very long period seismic signals. *Journal of Geophysical Research* 111: B06201.
Kumagai H and Chouet B (1999) The complex frequencies of long-period seismic events as probes of fluid composition beneath volcanoes. *Geophysical Journal International* 138: F7–F12.
Kumagai H and Chouet B (2000) Acoustic properties of a crack containing magmatic or hydrothermal fluids. *Journal of Geophysical Research* 105: 25493–25512.
Kumagai H, Chouet B, and Dawson PB (2005) Source process of a long-period event at Kilauea volcano, Hawaii. *Geophysical Journal International* 161: 243–254.
Kumagai H, Chouet B, and Nakano M (2002) Temporal evolution of a hydrothermal system in Kusatsu-Shirane Volcano, Japan, inferred from the complex frequencies of long-period events. *Journal of Geophysical Research* 107: ESE 9-1–ESE 9-10.
Kumagai H, Ohminato T, Nakano M, et al. (2001) Very-long-period seismic signals and caldera formation at Miyake Island, Japan. *Science* 293: 687–690.
Leet RC (1988) Saturated and subcooled hydrothermal boiling in groundwater flow channels as a source of harmonic tremor. *Journal of Geophysical Research* 93: 4835–4849.
Legrand D, Kaneshima S, and Kawakatsu H (2000) Moment tensor analysis of near-field broadband waveforms observed at Aso Volcano, Japan. *Journal of Volcanology and Geothermal Research* 101: 155–169.
Lokmer I, Saccorotti G, Lieto BD, and Bean CJ (2008) Temporal evolution of long-period seismicity at Etna Volcano, Italy, and its relationships with the 2004–2005 eruption. *Earth and Planetary Science Letters* 266: 205–220.
Margerin L (2013) Diffusion approximation with polarization and resonance effects for the modelling of seismic waves in strongly scattering small-scale media. *Geophysical Journal International* 192(1): 326–345.
McNutt SR (1996) Seismic monitoring and eruption forecasting of volcanoes: A review of the state-of-the-art and case histories. In: *Monitoring and Mitigation of Volcano Hazards*, pp. 99–146. New York: Springer-Verlag.
McNutt SR (2002) Volcano seismology. In: *International Handbook of Earthquake and Engineering Seismology*, pp. 383–406. Palo Alto, CA: IASPEI, ch. 25.
McNutt SR (2005) Volcanic seismology. *Annual Review of Earth and Planetary Sciences* 33: 461–491.
McNutt SR and Beavan RJ (1981) Volcanic earthquakes at Pavlof Volcano correlated with the solid earth tide. *Nature* 294: 518–615.
Metaxian J-P and Lasage P (1997) Permanent tremor of Masaya Volcano, Nicaragua: Wave field analysis and source location. *Journal of Geophysical Research* 102: 22529–22545.
Miller AD, Foulger GR, and Julian BR (1998) Non-double-couple earthquakes 2. Observations. *Reviews of Geophysics* 36: 551–568.
Minakami T (1974) Seismology of volcanoes in Japan. In: *Physical Volcanology*, pp. 1–27. Amsterdam: Elsevier.
Mogi K (1958) Relation between the eruptions of various volcanoes and the deformations of the ground surfaces around them. *Bulletin of the Earthquake Research Institute* 36: 99–134.
Molina I, Kumagai H, Garcia-Aristizabal A, Nakano M, and Mothes P (2008) Source process of very-long-period events accompanying long-period signals at Cotopaxi Volcano, Ecuador. *Journal of Volcanology and Geothermal Research* 176: 119.
Molina I, Kumagai H, and Yepes H (2004) Resonances of a volcanic conduit triggered by repetitive injections of an ash-laden gas. *Geophysical Research Letters* 31: L03603.
Mori J, Patia H, McKee C, et al. (1989) Seismicity associated with eruptive activity at Langila volcano, Papua New Guinea. *Journal of Volcanology and Geothermal Research* 38: 243–255.
Morita Y, Nakao S, and Hayashi Y (2006) A quantitative approach to the dike intrusion process inferred from a joint analysis of geodetic and seismological data for the 1998 earthquake swarm off the east coast of Izu Peninsula, central Japan. *Journal of Geophysical Research* 111: B06208.
Morrissey MM and Chouet B (1997) Burst conditions of explosive volcanic eruptions recorded on microbarographs. *Science* 275: 1290–1293.
Morrissey MM and Chouet B (2001) Trends in long-period seismicity related to magmatic fluid compositions. *Journal of Volcanology and Geothermal Research* 108: 265–281.
Nagaoka Y, Nishida K, Aoki Y, and Takeo M (2010) Temporal change of phase velocity beneath Mt. Asama, Japan, inferred from coda wave interferometry. *Geophysical Research Letters* 37(22) L22311.
Nakamichi H, Hamaguchi H, Tanaka S, Ueki S, Nishimura T, and Hasegawa A (2003) Source mechanisms of deep and intermediate-depth low-frequency earthquakes beneath Iwate volcano, northeastern Japan. *Geophysical Journal International* 154: 811–828.
Nakano M, Kumagai H, and Chouet B (2003) Source mechanism of long-period events at Kusatsu-Shirane Volcano, Japan, inferred from waveform inversion of the effective excitation functions. *Journal of Volcanology and Geothermal Research* 122: 149–164.

Nakano M, Kumagai H, Kumazawa M, Yamaoka K, and Chouet B (1998) The excitation and characteristic frequency of the long-period volcanic event: An approach based on an inhomogeneous autoregressive model of a linear dynamic system. *Journal of Geophysical Research* 103: 10031–10046.

Nettles M and Ekström G (1998) Faulting mechanism of anomalous earthquakes near Bardarbunga Volcano, Iceland. *Journal of Geophysical Research* 103: 17973–17984.

Neuberg J, Luckett R, Baptie B, and Olsen K (2000) Models of tremor and low-frequency earthquake swarms on Montserrat. *Journal of Volcanology and Geothermal Research* 101: 83–104.

Neuberg J, Luckett R, Ripepe M, and Braun T (1994) Highlights from a seismic broadband array on Stromboli volcano. *Geophysical Research Letters* 21: 749–752.

Nishida K, Kobayashi N, and Fukao Y (2000) Resonant oscillations between the solid Earth and the atmosphere. *Science* 287: 2244–2246.

Nishimura T (1998) Source mechanisms of volcanic explosion earthquakes: Single force and implosive sources. *Journal of Volcanology and Geothermal Research* 86: 97–106.

Nishimura T and Chouet B (2003) A numerical simulation of magma motion, crustal deformation, and seismic radiation associated with volcanic eruptions. *Geophysical Journal International* 153: 699–718.

Nishimura T, Hamaguchi H, and Ueki S (1995) Source mechanisms of volcanic tremor and low-frequency earthquakes associated with the 1988–89 eruptive activity of Mt. Tokachi, Hokkaido, Japan. *Geophysical Journal International* 121: 444–458.

Nishimura T, Nakamichi H, Tanaka S, et al. (2000) Source process of very long period seismic events associated with the 1998 activity of Iwate Volcano, northeastern Japan. *Journal of Geophysical Research* 105: 19135–19147.

Nishimura T, Ueki S, Yamawaki T, et al. (2003) Broadband seismic signals associated with the 2000 volcanic unrest of Mount Bandai, northeastern Japan. *Journal of Volcanology and Geothermal Research* 119: 51–59.

Ohminato T (2006) Characteristics and source modeling of broadband seismic signals associated with the hydrothermal system at Satsuma-Iwojima volcano, Japan. *Journal of Volcanology and Geothermal Research* 158: 467–490.

Ohminato T and Chouet B (1997) A free-surface boundary condition for including 3D topography in the finite-difference method. *Bulletin of Seismological Society of America* 87: 494–515.

Ohminato T, Chouet B, Dawson P, and Kedar S (1998) Waveform inversion of very long period impulsive signals associated with magmatic injection beneath Kilauea Volcano, Hawaii. *Journal of Geophysical Research* 103: 23839–23862.

Ohminato T and Ereditato D (1997) Broadband seismic observations at Satsuma-Iwojima volcano, Japan. *Geophysical Research Letters* 24: 2845–2848.

Ohminato T, Takeo M, Kumagai H, et al. (2006) Vulcanian eruptions with dominant single force components observed during the Asama 2004 volcanic activity in Japan. *Earth, Planets and Space* 58: 583–593.

Pitt AM and Hill DP (1994) Long-period earthquakes in the Long Valley caldera region, eastern California. *Geophysical Research Letters* 21: 1679–1682.

Prejean SG, Hill DP, Brodsky EE, et al. (2004) Remotely triggered seismicity on the United States west coast following the Mw 7.9 Denali fault earthquake. *Bulletin of Seismological Society of America* 94: S348–S359.

Pritchard ME, Rubin AM, and Wolfe CJ (2007) Do flexural stresses explain the mantle fault zone beneath Kilauea volcano? *Geophysical Journal International* 168: 419–430.

Ripepe M, Ciliberto S, and Schiava MD (2001) Time constraints for modeling source dynamics of volcanic explosions at Stromboli. *Journal of Geophysical Research* 106: 8713–8727.

Ripepe M and Gordeev E (1999) Gas bubble dynamics model for shallow volcanic tremor at Stromboli. *Journal of Geophysical Research* 104: 10639–10654.

Ripepe M, Marchetti E, Ulivieri G, et al. (2005) Effusive to explosive transition during the 2003 eruption of Stromboli volcano. *Geology* 33: 341–344.

Ripepe M, Poggi P, Braun T, and Gordeev E (1996) Infrasonic waves and volcanic tremor at Stromboli. *Geophysical Research Letters* 23: 181–184.

Rodgers PW (1968) The response of the horizontal pendulum seismometer to Rayleigh and Love waves, tilt, and free oscillations of the earth. *Bulletin of Seismological Society of America* 58: 1384–1406.

Ross A, Foulger GR, and Julian BR (1996) Non-double-couple earthquake mechanisms at The Geysers geothermal area, California. *Geophysical Research Letters* 23: 877–880.

Rowe CA, Aster RC, Kyle PR, Dibble RR, and Schlue JW (2000) Seismic and acoustic observations at Mount Erebus Volcano, Ross Island, Antarctica, 1994–1998. *Journal of Volcanology and Geothermal Research* 101: 105–128.

Rowe CA, Aster RC, Kyle PR, Schlue JW, and Dibble RR (1998) Broadband recording of Strombolian explosions and associated very-long-period seismic signals on Mount Erebus volcano, Ross Island, Antarctica. *Geophysical Research Letters* 25: 2297–2300.

Saccorotti G, Almendros J, Carmona E, Ibanez JM, and Pezzo ED (2001) Slowness anomalies from two dense seismic arrays at Deception Island volcano, Antarctica. *Bulletin of Seismological Society of America* 91: 561–571.

Saccorotti G, Chouet B, and Dawson P (2001) Wavefield properties of a shallow long-period event and tremor at Kilauea Volcano, Hawaii. *Journal of Volcanology and Geothermal Research* 109: 163–189.

Saccorotti G and Pezzo ED (2000) A probabilistic approach to the inversion of data from a seismic array and its application to volcanic signals. *Geophysical Journal International* 143: 249–261.

Sakai S, Yamada T, Ide S, et al. (2001) Magma migration from the point of view of seismic activity in the volcanism of Miyake-jima island in 2000. *Journal of Geography* 110: 145–155.

Sassa K (1935) Volcanic micro-tremors and eruption-earthquakes (Part 1 of the geophysical studies on the volcano Aso). *Memoirs of the College of Science, University of Kyoto* 18: 255–293.

Sassa K (1936) Micro-seismometric study on eruption of the Volcano Aso. *Memoirs of the College of Science, University of Kyoto* 19: 11–56.

Sato H, Fehler MC, and Maeda T (2012) *Seismic Wave Propagation and Scattering in the Heterogeneous Earth*, 2nd edn. Berlin: Springer-Verlag.

Scarpa R and Tilling RI (1996) *Monitoring and Mitigation of Volcanic Hazards*. New York: Springer-Verlag.

Schmidt RO (1986) Multiple emitter location and signal parameter estimation. *IEEE Transactions on Antennas and Propagation* 34: 276–280.

Schoenberg M and Sen PN (1983) Properties of a periodically stratified acoustic half-space and its relation to a Biot fluid. *Journal of the Acoustical Society of America* 73: 61–67.

Schuster GT, Yu J, Sheng J, and Rickett J (2004) Interferometric/daylight seismic imaging. *Geophysical Journal International* 157: 838–852.

Seidl D, Schick R, and Riuscetti M (1981) Volcanic tremors at Etna: A model for hydraulic origin. *Bulletin of Volcanology* 141: 43–56.

Sudo Y and Kong LSL (2001) Three-dimensional seismic velocity structure beneath Aso Volcano, Kyushu, Japan. *Bulletin of Volcanology* 63: 326–344.

Sutton AJ, Elias T, Gerlach TM, and Stokes JB (2001) Implications for eruptive processes as indicated by sulfur dioxide emissions from Kilauea Volcano, Hawaii, 1979–1997. *Journal of Volcanology and Geothermal Research* 108: 283–302.

Takagi N, Kaneshima S, Kawakatsu H, et al. (2006) Apparent migration of tremor source synchronized with the change in the tremor amplitude observed at Aso volcano, Japan. *Journal of Volcanology and Geothermal Research* 154: 181–200.

Takei Y and Kumazawa M (1994) Why have the single force and torque been excluded from seismic source models? *Geophysical Journal International* 118: 20–30.

Takeo M (1992) The rupture process of the 1989 offshore Ito earthquakes preceding a submarine volcanic eruption. *Journal of Geophysical Research* 97: 6613–6627.

Takeo M, Hamada N, Kashiwabara S, and Uhira K (1984) Analysis of long-period seismic waves excited by the explosive eruption of Mt. Asama on April 8. *Bulletin of the Volcanological Society of Japan* 29: 31–44.

Takeo M, Yamasato H, Furuya I, and Seino M (1990) Analysis of long-period seismic waves excited by the November 1987 eruption of Izu-Oshima volcano. *Journal of Geophysical Research* 95: 19377–19393.

Tameguri T, Iguchi M, and Ishihara K (2002) Mechanism of explosive eruptions from moment tensor analyses of explosion earthquakes at Sakurajima volcano, Japan. *Bulletin of the Volcanological Society of Japan* 47: 197–216.

Tsuruoka H, Kawakatsu H, and Urabe T (2009) GRiD MT (grid-based real-time determination of moment tensors) monitoring the long-period seismic wavefield. *Physics of the Earth and Planetary Interiors* 175: 8–16.

Tsutsui T and Sudo Y (2004) Seismic reflectors beneath the central cones of Aso Volcano, Kyushu, Japan. *Journal of Volcanology and Geothermal Research* 131: 33–58.

Uhira K, Baba T, Mori H, Katayama H, and Hamada N (2005) Earthquake swarms preceding the 2000 eruption of Miyakejima volcano, Japan. *Bulletin of Volcanology* 67: 219–230.

Uhira K, Ikeda S, and Takeo M (1994) Source process of explosion earthquakes deduced from short-period records at Sakurajima volcano. *Bulletin of the Volcanological Society of Japan* 40: 295–310.

Uhira K and Takeo M (1994) The source of explosive eruptions of Sakurajima volcano, Japan. *Journal of Geophysical Research* 99: 17775–17789.

Uhira K, Yamasato H, and Takeo M (1994) Source mechanism of seismic waves excited by pyroclastic flows observed at Unzen volcano, Japan. *Journal of Geophysical Research* 99: 17757–17773.

Ukawa M and Ohtake M (1987) A monochromatic earthquake suggesting deep-seated magmatic activity beneath the Izu-Ooshima Volcano, Japan. *Journal of Geophysical Research* 92: 12649–12663.

Vergniolle S, Brandeis G, and Mareschal J-C (1996) Strombolian explosions. 2. Eruption dynamics determined from acoustic measurements. *Journal of Geophysical Research* 101: 20449.

Waite GP, Chouet BA, and Dawson PB (2008) Eruption dynamics at Mount St. Helens imaged from broadband seismic waveforms: Interaction of the shallow magmatic and hydrothermal systems. *Journal of Geophysical Research* 113: B02305.

Wallace TC (1985) A reexamination of the moment tensor solutions of the 1980 Mammoth Lakes earthquakes. *Journal of Geophysical Research* 90: 11171–11176.

Watada S and Kanamori H (2010) Acoustic resonant oscillations between the atmosphere and the solid earth during the 1991 Mt. Pinatubo eruption. *Journal of Geophysical Research* 115: B12319.

West M, Sanchez JJ, and McNutt SR (2005) Periodically triggered seismicity at Mount Wrangell, Alaska, after the Sumatra earthquake. *Science* 308: 1144–1146.

White R (1996) Precursory deep long-period earthquakes at Mount Pinatubo: Spatiotemporal link to a basalt trigger. In: *Fire and Mud: Eruptions and Lahars of Mount Pinatubo Philippines*, pp. 307–327. Seattle, WA: University of Washington Press.

Widmer R and Zürn W (1992) Bichromatic excitation of long-period Rayleigh and air waves by the Mount Pinatubo and El Chichon volcanic eruptions. *Geophysical Research Letters* 19: 765–768.

Wielandt E and Forbriger T (1999) Near-field seismic displacement and tilt associated with the explosive activity of Stromboli. *Annali di Geofisika* 342: 407–416.

Wilson L and Head JW (1981) Ascent and eruption of basaltic magma on the Earth and Moon. *Journal of Geophysical Research* 86: 2971–3001.

Wolfe CJ, Okubo PG, and Shearer PM (2003) Mantle fault zone beneath Kilauea volcano, Hawaii. *Science* 300: 478–480.

Yamamoto M (2005) *Volcanic Fluid System Inferred from Broadband Seismic Signals*. PhD Thesis; University of Tokyo.

Yamamoto M and Kawakatsu H (2008) An efficient method to compute the dynamic response of a fluid-filled crack. *Geophysical Journal International* 174: 1174–1186.

Yamamoto M, Kawakatsu H, Kaneshima S, et al. (1999) Detection of a crack-like conduit beneath the active crater at Aso volcano, Japan. *Geophysical Research Letters* 26: 3677–3680.

Yamamoto M, Kawakatsu H, Yomogida K, and Koyama J (2002) Long-period (12 sec) volcanic tremor observed at Usu 2000 eruption: Seismological detection of a deep magma plumbing system. *Geophysical Research Letters* 29: 1329.

Yamamoto M and Sato H (2010) Multiple scattering and mode conversion revealed by an active seismic experiment at Asama volcano, Japan. *Journal of Geophysical Research* 115: B07304.

Yamamura K (1997) *Slow Waves in Solid-Liquid Two Phase System (in Japanese)*. Master's ThesisUniversity of Tokyo.

Yamamura K and Kawakatsu H (1998) Normal-mode solutions for radiation boundary conditions with an impedance contrast. *Geophysical Journal International* 134: 849–855.

Yamasato H, Fukui K, Uhira K, Hashimoto T, and Hori H (1993) Analysis of seismic and acoustic signals excited by pyroclastic flows at Unzen volcano. *Bulletin of the Volcanological Society of Japan* 38: 79–90.

This page intentionally left blank

4.16 Interaction of Solid Earth, Atmosphere, and Ionosphere

T Tanimoto, University of California, Santa Barbara, CA, USA
K Heki, Hokkaido University, Sapporo, Japan
J Artru-Lambin, Centre National d'Etudes Spatiales, Toulouse, France

© 2015 Elsevier B.V. All rights reserved.

This chapter is a revision of the previous edition chapter by T Tanimoto, J Artru-Lambin, Volume 4, pp. 421–444, © 2007, Elsevier B.V.

4.16.1	**Introduction**	422
4.16.2	**Seismic Noise**	423
4.16.2.1	Low-Frequency Peak: Atmospheric Effects	423
4.16.2.1.1	Cause	423
4.16.2.1.2	Turbulence	424
4.16.2.1.3	Reduction of seismic noise	424
4.16.2.2	Hum	425
4.16.2.2.1	Discovery	425
4.16.2.2.2	Seasonal variations	425
4.16.2.2.3	Excitation mechanism	425
4.16.2.2.4	Ubiquitous Rayleigh waves	426
4.16.2.3	Microseisms	426
4.16.2.3.1	Nature of waves	426
4.16.2.3.2	Excitation mechanisms	427
4.16.2.3.3	Source location	428
4.16.2.3.4	Microbaroms	428
4.16.2.3.5	Implication to past climate	429
4.16.2.4	From Noise to Structure	429
4.16.2.4.1	The correlation technique in diffuse wavefield	429
4.16.2.4.2	Related techniques	429
4.16.3	**Localized Sources of Interactions**	429
4.16.3.1	Historical Context	429
4.16.3.2	Theoretical Preliminaries	430
4.16.3.2.1	Wave propagation in the atmosphere	430
4.16.3.2.2	Frequency–wavelength domains of interest	431
4.16.3.3	Observation Techniques	431
4.16.3.3.1	Surface observation: Seismometer, microbarograph, and hydrophones	431
4.16.3.3.2	High-altitude observation: Ionosondes, Doppler sounding, transmission (GPS), and *in situ* measurements	432
4.16.3.4	Sources in the Solid Earth	433
4.16.3.4.1	Away from the source: Surface waves and tsunamis	434
4.16.3.4.2	Direct atmospheric waves	435
4.16.3.5	The Case of the Great Sumatra–Andaman Earthquake	436
4.16.3.6	The Cases of the 2011 Megathrust Earthquakes in Tohoku-Oki, Japan	437
4.16.3.7	Sources in the Atmosphere	438
4.16.3.7.1	Eruptions	438
4.16.3.7.2	Sonic boom	439
4.16.4	**Conclusion**	439
References		439

Nomenclature

C_s	Sound speed
e	Electron charge
f	Frequency in hertz (Hz)
f_1	Primary carrier frequency of GPS
f_2	Secondary carrier frequency of GPS
G	Gravitational constant (6.67×10^{-11} MKS unit)
g	Gravitational acceleration
k_x, k_y, k_z	Three components of wave number
m_e	Electron density
N_e	Electron density in the upper atmosphere
$\mathbf{v}$	Velocity (three-component vector)
z	Vertical coordinate ($z=0$ is surface and positive upward)
γ	Specific heat ratio
Δg	Perturbation to gravitational acceleration
ΔP	Surface pressure perturbation
$\Delta \rho$	Density perturbation
ζ	Ocean surface displacement

 http://dx.doi.org/10.1016/B978-0-444-53802-4.00083-X

$\bar{\zeta}$	Averaged ocean surface displacement over its wavelength	ω_a	Acoustic cutoff frequency
ρ	Density	ω_g	Brunt–Väisälä frequency
ω	$2\pi f$	ε_0	Permittivity of vacuum

4.16.1 Introduction

The atmosphere has been traditionally ignored by seismologists in their analysis of seismograms. Seismologists typically regard the surface of the solid Earth as the upper boundary of the Earth and this boundary is treated as a traction-free boundary. The atmospheric pressure is not zero, however, and the correct boundary condition must reflect the existence of the atmosphere with its diminishing density with height. Strictly speaking, the actual physical upper-boundary condition should be in the atmosphere as the radiation boundary conditions.

Despite this approximate treatment of the surface boundary, seismology has been successful mainly because of two factors. The first is that the impedance contrast at the atmosphere/solid Earth boundary is quite large. Density contrast across this layer is larger than 2000 (1.2 vs. 2700 kg m^{-3}) and P-wave velocity contrast is about 20 (300 vs. 6000 m s^{-1}). If we take the case of P-waves impinging on the solid surface from below, there is obviously some transmitted P-wave energy, but the majority of the energy is reflected back into the solid media due to this high impedance ratio. Resulting seismograms hardly indicate the effects of the leakage of transmitted energy.

The second important factor is the low mass of the atmosphere. The mass of the solid Earth is 6.0×10^{24} kg, while the mass of the atmosphere is only 5.1×10^{18} kg. Therefore, phenomena in the solid Earth are affected very little even if we include atmospheric effects in the analysis. It is thus justifiable to ignore the atmosphere and use the free-surface boundary conditions at the solid surface.

But does this mean that there are no benefits to include the atmospheric layer in the analysis? In a nutshell, the purpose of this chapter is to point out that there are many benefits to doing so. We can immediately think of two potential advantages; the first is the possibility of obtaining information on solid Earth processes from analyses of atmospheric waves. It has been pointed out that the eruption of Pinatubo volcano in 1991 generated air-coupled waves that provided information on the nature of the eruption process. Second, some earthquakes have been known to generate Rayleigh wave-coupled airwaves, which provide quantitative information on excitation source processes. These examples are analogous to the use of tsunami for the study of earthquake source process, because even though tsunami themselves are mostly confined to the liquid layer (ocean), their waveforms often provide information about the rupturing process on an earthquake fault. Analyses of atmospheric waves have not been as useful as analyses of tsunami so far, but more careful studies of atmospheric waves may lead to useful applications to understanding solid Earth processes.

Another benefit comes from the analysis of seismic noise. Most seismic noise in seismograms is generated by interactions between the solid Earth and the atmosphere or the ocean. In the past 10 years, such noise was shown to be practically useful because we can recover Green's functions between any pair of stations using a cross correlation technique. Its application to Earth structure study is under way: Seismic noise can provide surface wave dispersion curves between two stations and promise to provide results that were not available from earthquake data. The cross correlation technique also provides information on the flow (direction) of propagating seismic energy, thereby providing potential information about the source location of noise, which was not possible to determine before.

This chapter consists of two main parts: In the first part, we will describe the nature of seismic noise and Green's function cross correlation technique. We will review the study of seismic noise and how noise illuminates our understanding of atmosphere–solid Earth interactions. Since the atmosphere and oceans are strongly coupled and often not separable, the oceans are involved in many of these processes.

In the second part, we will discuss the effects of transient sources in the solid Earth such as earthquakes, volcanic eruptions and large atmospheric explosions, and their resulting effects in the atmosphere. These events radiate energy in various types of waves, such as seismic waves in the solid Earth, tsunamis in the oceans, and internal waves in the atmosphere. Dynamic coupling allows the transmission of a part of the wave energy across the Earth surface, hence the generation of atmospheric, ocean, or tsunami waves from sources external to their respective domains. Sources of such coupled waves include (1) atmospheric explosions, natural (meteors and volcanoes) or man-made, (2) earthquakes, and (3) sonic booms from supersonic jets (Concorde and Space Shuttle). Some notable effects observed after the great Sumatra–Andaman earthquake will also be discussed.

The first and second parts take two different views on modeling the source–medium relationship. The first part follows the view that seismic waves generated by atmospheric effects are generated by surface atmospheric pressure variations. The Earth is basically separated into two distinct media in contact. In this view, generation of acoustic/gravity waves in the atmosphere by sources in the solid Earth is mainly by vertical surface displacement. On the other hand, the second follows the so-called full-coupling approach; there is no artificial boundary between the atmosphere and the solid Earth, and the whole Earth is treated as a single medium. Normal modes are the normal modes of the whole system, and propagating waves cross the atmosphere–solid Earth boundary just like other boundaries inside the Earth. The approach in the first part is an approximation in comparison with a more holistic approach in the second part, but it seems to hold well and tend to provide better physical insights.

4.16.2 Seismic Noise

Peterson (1993) summarized seismic noise characteristics from global distribution of 75 stations. One of the most frequently quoted results from this study is the parameterized model called the new low-noise model (NLNM). Figure 1 shows a plot of this model in the form of acceleration power spectral density. This model has often been used as a reference in the discussion of seismic noise studies.

From the perspective of a seismologist, one of the main features in Figure 1 is the existence of a low-noise frequency band from 2 to 20 mHz (millihertz). Existence of this low-noise band explains why analyses of long-period body and surface waves have been successful in the past few decades as these waves exist in this low-noise band (1 mHz–10 Hz). Detection of small amplitude waves has been possible because of this low-noise characteristic, which is not as easily done outside this frequency band.

From the perspective of a scientist interested in the atmosphere–solid Earth interactions, Figure 1 shows three prominent peaks; from the low-frequency end, they are (1) the low-frequency peak below 0.1 mHz, whose peak is outside the range of this plot but the decreasing trend with frequency up to about 2 mHz is clearly recognized, (2) a small peak or a bump at about 7–10 mHz (denoted as hum in this figure), and (3) large amplitude peak(s) at 0.1–0.2 Hz (microseisms). In this section, we will discuss each of the three peaks from the low-frequency end, particularly focusing on the question which components in the Earth system interact to create these peaks.

The model NLNM was derived from vertical component seismograms. The noise in horizontal components is often an order of magnitude larger than the noise in vertical component because horizontal components are affected by tilts that are generated by temperature variations and atmospheric pressure variations. They are often dominant sources of noise in horizontal component seismographs but are confined to shallow depths and local regions. In this chapter, noise in horizontal components will not be discussed.

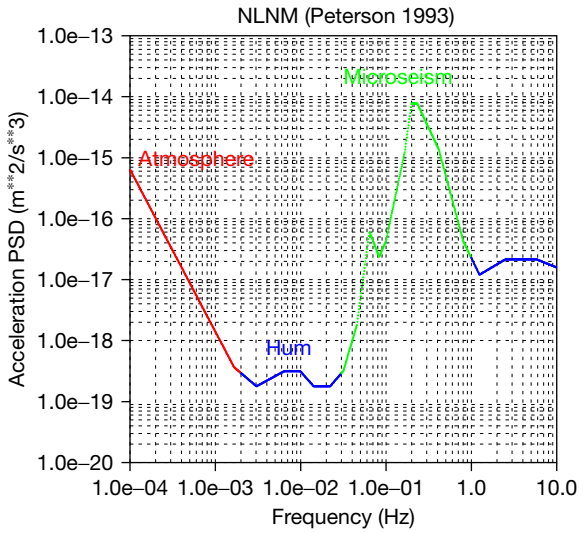


Figure 1 The new low-noise model (NLNM) by Peterson (1993). Power spectral density (PSD) in unit for acceleration is used for this plot. Frequency range is from 1/10 000 to 10 Hz.

4.16.2.1 Low-Frequency Peak: Atmospheric Effects

Seismic noise below 2 mHz increases toward lower frequencies (Figure 1). This is a global feature that has been confirmed repeatedly from long-period seismographs and gravimeters.

4.16.2.1.1 Cause

The main cause of this noise is density changes in the atmosphere. For example, as weather patterns change, density changes occur in the atmosphere. These changes affect the gravitational field and generate small perturbations in gravitational acceleration at the surface of the Earth.

This mechanism was pointed out by Warburton and Goodkind (1977) by analyzing two colocated instruments, a barometer and a superconducting gravimeter. Barometric (surface pressure) data represent an integration of density anomalies along a vertical column above a station:

$$\Delta P = \int_0^\infty \Delta\rho g \, dz \qquad [1]$$

where ΔP is the surface pressure perturbation, $\Delta\rho$ is density perturbation in this formula, and the integration is from the surface of the Earth ($z=0$) to infinity. Effects of atmospheric density perturbations above an observing station can be approximated by a formula similar to the Bouguer gravity formula:

$$\Delta g = 2\pi G \int_0^\infty \Delta\rho \, dz \qquad [2]$$

where Δg is the perturbation to gravitational acceleration and G is the gravitational constant. Since g does not vary very much near the surface, barometric data and gravity data are related, to a good approximation, by

$$\Delta g = \frac{2\pi G}{g} \Delta P \qquad [3]$$

Warburton and Goodkind showed that colocated gravimeter data and barometric data correlate with coefficients close to a value predicted by $2\pi G/g$.

There is, however, an assumption in using the Bouguer gravity formula; in applying this formula, density perturbation in the atmosphere is assumed to be shaped like a disklike pattern, and density variations laterally away from the station location (in latitude and longitude) do not affect the gravitational acceleration very much. More careful evaluation of such effects, using a meteorologic simulation model, showed that deviation from this formula is small (Boy et al., 1998; Hinderer and Crossley, 2000).

Changes in the atmospheric pressure also cause deformation of the elastic Earth as a surface load. Correction due to this effect is usually added to the aforementioned formula in the analyses (Crossley et al., 1995; Merriam, 1992; Niebauer, 1998; Spratt, 1982; Van Dam and Wahr, 1987; Warburton and Goodkind, 1977). This effect is not negligible and in fact was claimed to be larger than the previously mentioned effect in earlier period of seismic noise study (Sorrells, 1971; Sorrells et al., 1971). But it is smaller than the right-hand side of the aforementioned formula.

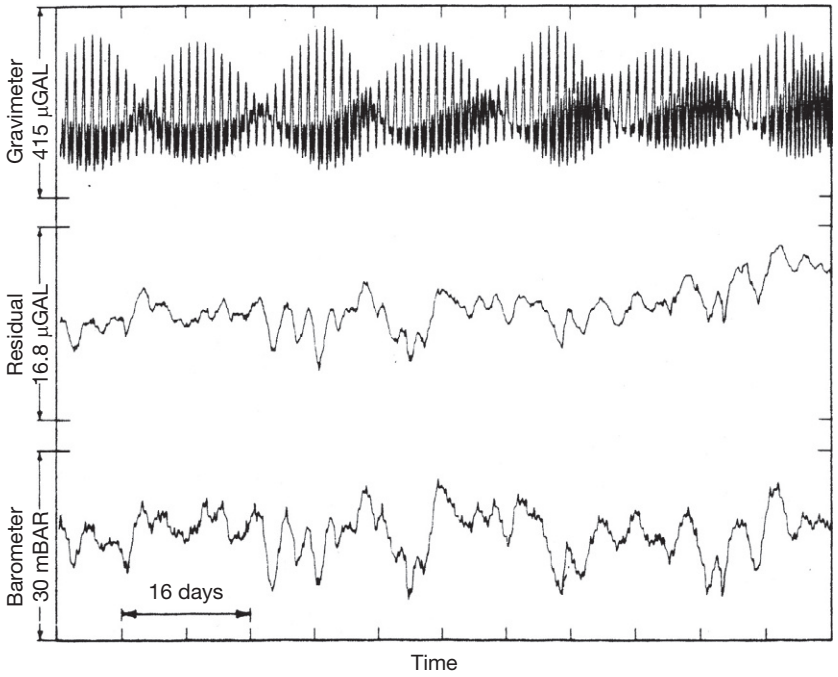


Figure 2 (Top) Original gravity signal from Warburton and Goodkind (1977). Data were recorded by superconducting gravimeter. (Middle) Gravity signal after tidal signal removal. (Bottom) Barometer signal, which correlates well with the trace in the middle.

A clear demonstration of atmospheric pressure effects at the time of a cold front passage was documented by Müller and Zürn (1983), and some careful quantitative computational analysis was reported by Rabbel and Zschau (1985).

Figure 2 shows the evidence of correlation between gravity data and surface pressure data after the removal of tidal signals from gravimeter data. The original gravity record (top) is dominated by tidal signals, but after their removal, the middle trace is obtained. This time series matches the barometric data at the bottom closely, indicating peak-to-peak correlations between the middle trace and the bottom trace.

4.16.2.1.2 Turbulence

While it seems that the correlation between gravity (ground motions) and surface pressure data is good as shown in Figure 2, the correlation coefficient hovers around 0.5 and never becomes close to one. The uncorrelated part of the signal may be due to additional forces, such as wind stress (Reynolds stress) in atmospheric turbulence, which dominates the frequency range above 1 mHz (Tanimoto, 1999). Frequency dependence of surface pressure in this frequency band shows the $1/f$ behavior as observed spectra in Figure 3 show (e.g., Tanimoto and Um, 1999). This $1/f$ behavior is consistent with the prediction by the theory of turbulence with the Kolmogorov-type scaling relation (Landau and Lifshitz, 1987; Tennekes and Lumley, 1971) and directly supports the importance of turbulence in this frequency band.

4.16.2.1.3 Reduction of seismic noise

Existence of the previously mentioned mechanism suggests that, if a seismic/gravity instrument is colocated with a barometer, one can reduce the level of noise considerably by taking advantage of this correlation. For analyses of tidal and lower-frequency data, this has become a common practice since the study of Warburton and Goodkind (1977). For seismic data, especially for normal-mode studies (about 1 mHz and above), Zürn and Widmer (1995), Beauduin et al. (1996), and Roult and Crawford (2000) showed significant noise reduction. These studies showed that noise level below 1 mHz can be made lower than the model NLNM by this correlation technique.

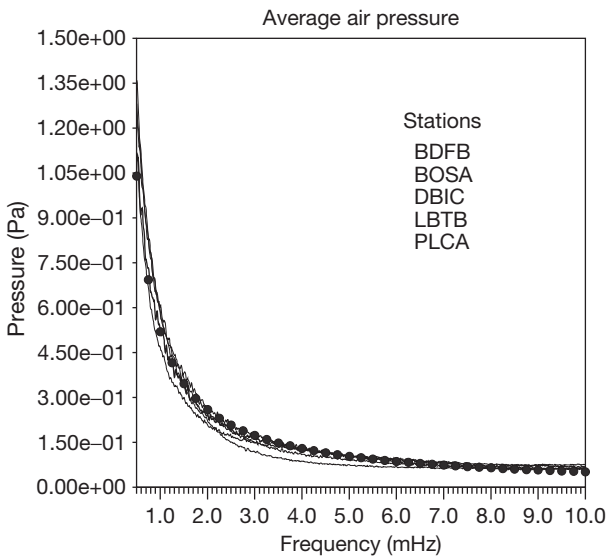


Figure 3 Barometric data (pressure) from five stations in North and South America. Analysis was performed for 20-day-long signals and spectral amplitudes were averaged over 0.2 mHz range. Solid circles denote $1/f$ trend.

4.16.2.2 Hum

The second peak denoted by hum in Figure 1 was identified as a broad peak by Peterson (1993). It was shown later that this peak is associated with multiple modal peaks, primarily on the left-hand side of its maximum. Figure 4 shows an example from data in this frequency range, obtained from global network data by averaging 11 stations located at various parts of the world. The bottom panel is an enlarged figure within the small box in the top panel and shows that modal peaks exist for the frequency range between 2 and 7 mHz. They are shown to match the eigenfrequencies of fundamental spheroidal modes almost exactly, as the eigenfrequencies for the preliminary reference Earth model (PREM; Dziewonski and Anderson, 1981) are drawn in the bottom panel by vertical lines. These continuously excited modes are all fundamental spheroidal modes and do not seem to contain any overtone modes. They are now commonly referred to as the Earth's hum.

4.16.2.2.1 Discovery

The discovery of continuously excited modal peaks was made in 1997. A broad peak in Peterson's model has been noted since 1993, or perhaps even earlier, but the hum was discovered as an independent feature from it. The initial reports were by Suda et al. (1998), Kobayashi and Nishida (1998), and Tanimoto et al. (1998). These and subsequent studies showed that all sites where the noise level is as low as 10^{-18} ($m^2\ s^{-3}$) in acceleration spectral density (for the frequency band 2–15 mHz) show signals of the hum. Since the global minimum of horizontal noise is an order of magnitude higher than this value, all observations were from vertical component seismograms or gravimeters. The types of seismic instruments that led to these discoveries included broadband seismometers STS-1 and gravimeters with spring sensor (LaCoste–Romberg gravimeters) or superconducting sensors. It was recently shown that seismic data from broadband seismometers STS-2 (with lower pendulum period than STS-1, about 100 s) at the Black Forest Observatory (BFO) in Germany show the hum (Widmer-Schnidrig, 2003). This may not be surprising, however, since the noise level is as low as 10^{-18} ($m^2\ s^{-3}$) for STS-2 at this low-noise site. The most important point seems to be that the signal of the hum is observed if the noise level is about 10^{-18} ($m^2\ s^{-3}$) or less, in unit for acceleration spectral density.

Nawa et al. (1998) reported the existence of the hum in their superconducting gravimeter data at Showa Station in Antarctica. Their report was made before the three papers cited earlier, but the reported results contain peaks from grave spheroidal modes (angular degree <10) that have not been confirmed independently. In addition, their later study (Nawa et al., 2000a,b) showed that the noise level at this station seems to be much higher than 10^{-18} ($m^2\ s^{-3}$) for the frequency band 1–5 mHz.

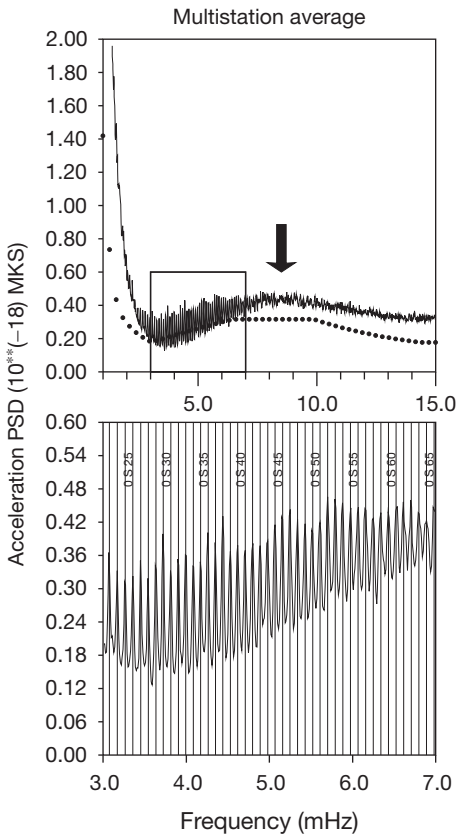


Figure 4 Averaged acceleration PSD from 11 stations, distributed from various parts of the world, is shown at top. Circles in the top panel are NLNM. Spectra in the small box (top) are enlarged in the bottom panel. Each modal peak is shown to match the eigenfrequency of the fundamental spheroidal mode (vertical lines in the bottom panel).

4.16.2.2.2 Seasonal variations

These modal amplitudes display seasonal variations; claims of the predominant 6-month periodicity were made by Tanimoto and Um (1999) from a frequency-domain analysis and by Ekström (2001) from an entirely independent time-domain analysis. A claim for an annual seasonality was made by Nishida et al. (2000), although they claimed that annual signal was dominant and was thus mildly different from the previously mentioned two studies.

Detection of seasonal variations had major implications for the source of excitation of the hum because it basically removed causes in the solid Earth, because phenomena in the solid Earth do not usually have clear seasonal signatures. Up until these discoveries on seasonality were made, slow earthquakes (e.g., Beroza and Jordan, 1990) were considered to be one of the major candidates for the cause of the hum.

4.16.2.2.3 Excitation mechanism

The cause of the hum may be in the atmosphere or in the oceans. The atmospheric excitation was advocated by Kobayashi and Nishida (1998), Tanimoto and Um (1999), and Fukao et al. (2002). In these papers, pressure fluctuations in the turbulent atmosphere were postulated as the cause. Modal amplitudes of individual peaks were shown to be explained by the atmospheric excitation model. But there was an uncertain part in this scenario, particularly on the correlation length in atmospheric turbulence within the frequency band (2–7 mHz range). This was critical because the excitation of modes by turbulent atmosphere is proportional to this parameter (Goldreich and Keeley, 1977; Tanimoto, 1999). The atmospheric excitation mechanism was advocated by assuming that this correlation length was about 1 km or larger

for the frequency band 2–7 mHz. This correlation length has not been confirmed by observation, although the existence of turbulent boundary layer of thickness 1 km seems to imply that it may be a viable candidate.

In addition to this uncertainty in source strength, the atmospheric excitation hypothesis does not provide a good reason for the existence of the broad noise peak between 3 and 15 mHz. In both Tanimoto and Um (1999) and Fukao et al. (2002), this broad spectral peak was assumed to be caused by an unknown background noise, and only the modal peaks above this background noise were modeled by the atmospheric pressure variations at the surface.

The oceanic excitation hypothesis was advanced by Rhie and Romanowicz (2004) and Tanimoto (2005). Rhie and Romanowicz (2004) used seismic array data and located the sources of Rayleigh waves (noise) using two arrays in Japan and California. They found the excitation sources to be in the oceans, especially in the midlatitude bands (30°–60°) in the northern and southern hemispheres. They also claimed that source locations switched rather abruptly between the northern and southern hemispheres. This feature seems to be compatible with general patterns of storm behaviors, although further independent confirmation is desirable.

Tanimoto (2005) showed that the overall spectral shape of the hum, the modal peaks, and the broad spectral peak depicted in **Figure 1** (3–15 mHz) can be explained by a single mechanism if the oceanic infragravity waves were the cause. A rather ad hoc feature in the atmospheric excitation hypothesis, which has to find separate causes for modal peaks and for the broad spectral peak, can then be avoided. Nishida et al. (2005a,b) argued that it may still be possible to create this broad peak by atmospheric effects, but a detailed mechanism is still missing in the atmospheric excitation hypothesis.

One of the uncertainties in the oceanic excitation hypothesis lies in our lack of knowledge on the oceanic infragravity waves. A limited number of observations (e.g., Watada and Masters, 2001) are now supplemented by new observations. Also, some new understanding as to the generation of infragravity waves from oceanic swells is emerging from observation (Dolenc et al., 2008). Although there have been some studies on the oceanic infragravity waves (Bromirski and Gerstoft, 2009; Dolenc et al., 2008; Okihiro et al., 1992; Traer et al., 2012; Webb et al., 1991) and many results point to near-coastal generation of waves, a comprehensible, total picture of the generation mechanism seems to be missing. This is still an unresolved question.

Satellite ocean-wave data provide semihemispheric switching of activities, as shown in **Figure 5**. Ocean waves occasionally reach 10 m or more in high-activity regions, which generates pressure perturbations higher than surface atmospheric pressure. This behavior seems to explain the 6-month periodicity naturally as well as source locations of Rayleigh waves observed by Rhie and Romanowicz (2004). This satellite evidence does not necessarily prove the oceanic excitation mechanism for the hum because strong atmospheric winds are associated with these ocean-wave behaviors and the atmosphere also contains 6-month periodicity. There is clearly an inherent difficulty in the argument of atmospheric versus oceanic excitation because the atmosphere and the oceans are coupled in almost all scales.

It is most likely that the next breakthrough will come from an improved understanding of ocean gravity waves (e.g., Bromirski and Stephen, 2012; Bromirski et al., 2010; Uchiyama and McWilliams, 2008), although it may take some years.

4.16.2.2.4 Ubiquitous Rayleigh waves

One of the most important notions developed in these studies is that the energy associated with the hum, for the entire frequency range 3–15 mHz, consists of Rayleigh wave energy. From observation, Nishida et al. (2002) showed that signals between 2 and 20 mHz have similar phase velocities to Rayleigh wave phase velocities predicted by the PREM (Dziewonski and Anderson, 1981). Nishida et al. (2009) performed tomographic inversion from surface wave data retrieved by an application of the cross correlation technique to noise data. The results confirmed consistency with previously derived tomographic results from earthquake data. Tanimoto (2005) showed that the whole spectra in this frequency band can be synthesized by normal-mode summation of spheroidal modes, potentially excited by the oceanic infragravity waves, thereby indicating that they are Rayleigh waves.

While Love waves seem to share some fraction of seismic noise, our understanding of it lags as it is hard to analyze horizontal component data.

4.16.2.3 Microseisms

It has been noted since the early twentieth century that the most obvious and perhaps annoying noise in seismograms is the microseisms with the peak frequency at about 0.1–0.4 Hz (**Figure 1**). Amplitudes of this noise were so overwhelming that, before the development of high-dynamic-range digital seismic instruments, seismologists recorded seismic waves separately for high-frequency range (above about 0.5 Hz) and for low-frequency range (below 0.1 Hz) in order to avoid this microseismic noise. World-Wide Standard Seismograph Network, which played the central role for the development of global seismology from the 1960s to 1980s, had such separate (short-period and long-period) instruments.

Modern instruments with high-dynamic-range and digital recordings removed such a cumbersome recording procedure (Wielandt and Steim, 1986). They record seismic signals from about 1 mHz to 10 Hz using the same sensor. A benefit of such recording procedure is that microseisms are now recorded continuously and this has raised interests among some seismologists since they can now analyze microseism data based on modern high-dynamic-range digital data.

4.16.2.3.1 Nature of waves

A large fraction of secondary microseisms consist of Rayleigh waves with smaller fraction of Love wave energy. This was shown by Lee (1935) and Haubrich et al. (1963) by particle motion analysis, which indicated a characteristic retrograde elliptical particle motion. Later, array analysis (Capon, 1972; Lacoss et al., 1969) showed that some higher-mode energy and Love wave energy were mixed in the signals. Earlier, Gutenberg (1958) also discussed two types of microseisms, apparently referring to predominant Rayleigh waves and occasional S-waves. From modern three-component data, it is easy to confirm the dominance of Rayleigh waves in microseisms

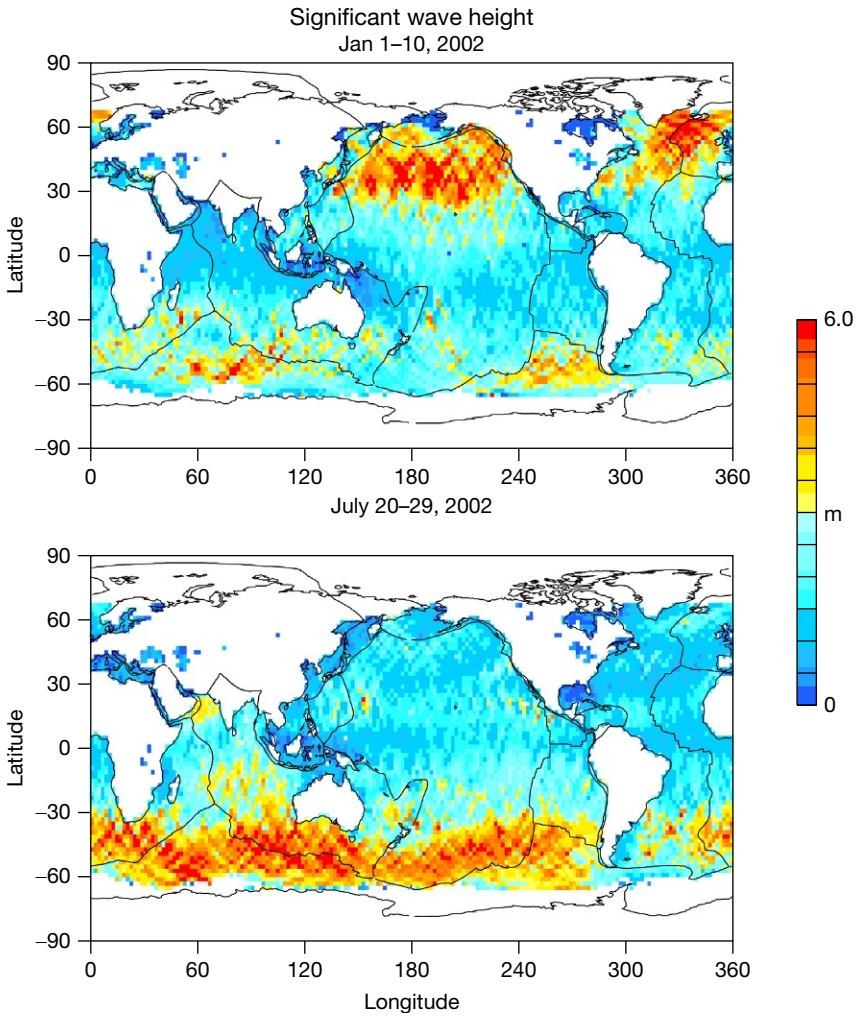


Figure 5 Significant wave height data from TOPEX/Poseidon. Data from two selected periods, one in January (top) and the other in July (bottom), are shown. In January, ocean waves are energetic in the midlatitude region in the northern hemisphere, while ocean waves in the southern hemisphere become energetic in July.

from phase velocity measurements from array observations (Capon, 1972; Lacoss et al., 1969) or from phase-shift observations between horizontal and vertical components, which indicate mostly 90° phase shifts.

Primary microseisms, whose frequency band is about 0.07 Hz, seem to contain a larger fraction of Love waves. This seems to imply different excitation mechanisms between the primary and the secondary microseisms. The commonly accepted mechanism for the secondary microseisms, the wave–wave interaction, is basically a vertical forcing (Longuet-Higgins, 1950), which excited Rayleigh waves much better than Love waves. The mechanism for the primary microseisms is more likely to be interactions of ocean waves with the solid Earth in shallow oceans and thus contains horizontal forcing through topographic coupling (Saito, 2010).

4.16.2.3.2 Excitation mechanisms

One of the important characteristics in microseisms is the fact that larger amplitudes occur for the double-frequency microseisms (0.1–0.4 Hz) than the primary-frequency microseisms that have the same predominant frequencies with ocean swells (0.05–0.07 Hz). It is still not clear why the double-frequency microseisms have such large amplitudes, but the most widely accepted (basic) mechanism for the double-frequency microseisms is that of Longuet-Higgins (1950). Longuet-Higgins (1950) showed that interactions of two ocean waves in opposite direction can create the double-frequency microseisms through the nonlinear (advection) term in the Navier–Stokes equation. The crux of his theory is that, for surface displacement ζ, pressure perturbation at sea bottom is given by

$$p(t) = \rho \frac{\partial^2}{\partial t^2} \left\{ \frac{1}{2} \zeta^{-2} \right\} \qquad [4]$$

even when the depth extent of colliding waves, which make surface displacement ζ, does not reach the sea bottom. In fact, this pressure arises as a constant in the Bernoulli equation (Longuet-Higgins, 1950), implying that this pressure occurs at all depths. The bar denotes an averaging procedure over a wavelength. For example, if two ocean waves with the same frequency are propagating in opposite directions,

$$\zeta(x, t) = a_1 \cos(\omega t - kx) + a_2 \cos(\omega t + kx) \quad [5]$$

we get

$$p(t) = -\rho a_1 a_2 \omega^2 \cos(2\omega t) \quad [6]$$

from Eqn [4]. This formula contains two important features; first, it shows the occurrence of double-frequency pressure variations. These double-frequency pressure variations at sea bottom generate the double-frequency microseisms. Second, this formula shows that if there exists a unidirectional propagating wave, say, $+x$ meaning $a_2 = 0$, this pressure term goes to zero because it is proportional to $a_1 a_2$. The double-frequency pressure variation requires existence of colliding waves. This mechanism was experimentally confirmed by Cooper and Longuet-Higgins (1951). Phillips (1977) showed a derivation of the same formula from the Navier–Stokes equation by carefully analyzing the vertical momentum balance.

It should be noted, however, that this derivation was made for an incompressible fluid. Compressibility of the ocean is important for the dominant, secondary microseisms (Longuet-Higgins, 1950).

Hasselmann (1963) cast the problem in a more general context, expressing the ocean wave field as a wave number integral and then showing that the predominant term is equivalent to the Longuet-Higgins formula. Normal-mode excitation analysis of seismic wavefields by nonlinear interactions of ocean waves also showed recently that the Longuet-Higgins pressure formula naturally arises from the analysis (Tanimoto, 2007a).

The excitation mechanism for the primary microseisms should be different from the secondary microseisms. Their characteristics are the following: (i) they have the same frequency with ocean waves and (ii) there seems to be a larger fraction of Love waves than secondary microseisms (Friedrich et al., 1998). These characteristics suggest that the excitation is more likely to be through the interactions of ocean waves with the solid Earth at sea bottom, for example, through topographic couplings (Saito, 2010).

4.16.2.3.3 Source location

While it is clear that the nonlinear interactions of ocean waves are necessary to generate microseisms, there is some confusion as to exact locations of excitation sources. There has been a long debate about near-coastal sources versus pelagic sources (in the deep oceans). Near-coastal sources seem common but there seems to be a strong case for some particular pelagic sources (e.g., Cessaro, 1994; Kedar et al., 2008). What we know for sure is the fact that a source must be in a place where ocean-wave collisions occur.

Close relationship between microseisms and ocean swells near the shore was shown by Haubrich et al. (1963), Bromirski et al. (1999), Cessaro (1994), Gerstoft and Tanimoto (2007), Tanimoto (2007b), and Traer et al. (2012). Bromirski et al. (1999) and Tanimoto (2007b) examined a correlation between microseisms from seismograms and ocean waves from buoy data and demonstrated a high correlation and thus a causal relationship. Cessaro (1994) and Gerstoft and Tanimoto (2007) used array analysis from three separate networks and located the sources of microseisms.

It was often conjectured in early studies of microseisms that storms provide a condition that ocean waves collide from opposite directions near their eyes. If this were true, it could be an example of a pelagic source. Tracking of a storm was done by Santo (1960), Sutton and Barstow (1996), and Bromirski (2001) in order to understand the source process of the double-frequency microseisms. The work by Bromirski (2001) provided the best data set and clearly showed that the dominant source area for the double-frequency microseisms was not in the open ocean, where the highest waves occurred, but was near the coast. It seems therefore that standing ocean waves that occur near the coasts, even at the time of large low-pressure system, are the source of double-frequency microseisms. A similar example was shown by Friedrich et al. (1998). This study performed a careful array analysis of seismic data and reported that the primary microseisms contained more Love waves than the double-frequency microseisms, suggesting that the exciting mechanisms may be quite different between them.

Kedar et al. (2008) presented a fairly strong case for a pelagic source in the North Atlantic Ocean. This was based on a comparison between seismic data and hindcase ocean-wave spectral data from the North Atlantic Ocean. They also claimed that there is a preferential ocean depth range, based on the Longuet-Higgins (1950) theory. The source location of this study is close to one of the two areas, pointed out by Cessaro (1994). The other location was in the northeast Pacific, close to Canada and Alaskan coasts. Other recent studies suggested the existence of pelagic sources are Ardhuin et al. (2011), Hillers et al. (2012), and Zhang et al. (2010).

Schulte-Pelkum et al. (2004) showed a temporary effect of microseismic signals in California caused by ocean waves from the Atlantic Ocean, an evidence for a pelagic source, although occurrence of such phenomena does not seem to be common. Their Atlantic source may have been identical with the source found by Kedar et al. (2008) as this pelagic source seems persistent.

It has also been noted that source area of primary microseisms is often spread out along the coasts, whereas source areas of the double-frequency microseisms appear to be specific locations. This was observed in Europe (Friedrich et al., 1998), on the Atlantic coast (Bromirski, 2001), and on the Pacific coast of the United States (Bromirski et al., 1999; Schulte-Pelkum et al., 2004).

4.16.2.3.4 Microbaroms

Microbaroms are atmospheric low-frequency waves, having frequencies close to those of double-frequency microseisms. The similarity of power spectra to those of microseisms suggests that they are of the same origin (e.g., Donn and Naini, 1973; Nishida et al., 2005a,b). Posmentier (1967) presented a theory analogous to the mechanism proposed by Longuet-Higgins (1950). Arendt and Fritts (2000) published a more complete theory, carefully analyzing various types of waves that arise from interactions of ocean waves. A more elaborate derivation was given by Waxler and Gilbert (2006) who analyzed a two-fluid model (air over seawater) for the radiation of acoustic energy by ocean waves. The basic mechanism in these studies is the same with the Longuet-Higgins mechanism for microseism generation and assumes

interactions of two surface waves propagating in opposite directions. Standing ocean waves near the coasts appear to be generating both microseisms and microbaroms.

4.16.2.3.5 Implication to past climate

Changes in the climate are likely to be related to ocean-wave behaviors (Bromirski and Duennebier, 2002; Bromirski et al., 1999), which may threaten coastal areas if they become too energetic. However, recovery of past ocean-wave heights from meteorologic data is generally difficult. Grevemeyer et al. (2000) proposed to use activities of microseisms in historical seismograms as a proxy for obtaining information for historical ocean-wave behavior. They reported that the number of high microseismic days increased from 7 to 14 days in the last 50 years and suggested that this may be related to global warming indirectly. While this result should be regarded as preliminary, it is reasonable to expect some correlation, and microseism records may turn out to be a useful source of information for ocean-wave behavior and thus for climate changes.

4.16.2.4 From Noise to Structure

One of the current motivations to study seismic noise is its potential use for Earth structure study. Many techniques have been developed, each focusing on different aspects of data; examples include, in historical order, spatial variation of correlation amplitudes, array analysis for Rayleigh wave phase velocity measurements, use of horizontal–vertical amplitude ratios for Rayleigh wave signals, and the two-station correlation technique to recover Green's functions between a pair of stations. Later, we first discuss the most recent development, the correlation technique to recover Green's functions in diffuse seismic wavefields, and then discuss other techniques under classification as the traditional methods.

4.16.2.4.1 The correlation technique in diffuse wavefield

The first application of the correlation technique to recover structural information was demonstrated in helioseismology (Duvall et al., 1993). The essence of the technique was to recover travel time versus distance curve for multiply reflected waves through cross correlation of data at various distances. In the case of the Sun, seismic sources are stochastic in time and space, and thus, there are no specific noise sources; it can be anywhere and anytime. Therefore, such a technique was most needed and worked effectively to get the results.

Applications to other fields, such as geophysical exploration (Rickett and Claerbout, 1999) and ultrasonics (Weaver and Lobkis, 2001), followed before seismology adopted it recently (Campillo and Paul, 2003). Theoretical basis to recover Green's functions from noise was discussed in Weaver and Lobkis (2001), Lobkis and Weaver (2001), van Tiggelen (2003), Snieder (2004), and Roux et al. (2005a).

Applications to seismic data started with Campillo and Paul (2003). Recovered signals were dominated near the microseismic frequency bands (about period 5–10 s) in earlier demonstrations. The technique was soon applied to local- to regional-scale problems such as Southern California (Sabra et al., 2005a,b; Shapiro et al., 2005). It is now applied to much wider frequency bands (up to 100 s in period) and is also applied to continental-scale structures.

Typical seismic applications first recover Green's functions by cross correlation of seismograms at two stations. Since data are mainly from vertical components and are dominated by Rayleigh waves, traditional dispersion measurements for group velocity are often used to retrieve Earth structure. Dominance of fundamental-mode surface waves in noise is perhaps unavoidable because the excitation sources are in the atmosphere and the oceans, although there are now some studies that report successful recovery of body wave signals (e.g., Roux et al., 2005b). This is currently a rapidly expanding field and the landscape of the research field is expected to change quickly in a few years.

4.16.2.4.2 Related techniques

Aki (1957, 1965) proposed a method, often referred to as the spatial autocorrelation method (SPAC), to determine the local seismic structure. This technique also uses the cross correlation among stations but specifically uses a coherency between two stations; in a flat layered media, the real part of coherency becomes proportional to a Bessel function (Aki, 1965) from which one can obtain phase velocity. Extension of this approach to an attenuating medium was attempted by Prieto et al. (2009), and some theoretical important aspects of the approach were clarified by Tsai (2011).

Lacoss et al. (1969) showed that an array-based phase velocity measurement is a powerful approach. This is basically a beamforming technique, using array data, and is becoming increasingly popular because of availability of dense seismographic networks.

Both SPAC and the beamforming approach are still used at present and an enormous body of literature exists in this area, especially in geotechnical engineering. In geotechnical engineering, the frequency range extends up to 10–20 Hz, exceeding the microseismic frequency band (0.1–0.4 Hz). These studies use not only microseismic signals but also higher-frequency signals generated by other sources and can only be called a cultural noise. A good recent summary of this field was given by Okada (2003).

These traditional techniques have been around 40–50 years, but their practical use may expand as quality and density of seismic instruments have improved recently. We expect to see more use of these methods because there are urgent needs, especially in urban area, to understand near-surface structure as shallow seismic structure plays a critical role for ground motion amplification at the time of major earthquakes.

4.16.3 Localized Sources of Interactions

In this section, we will focus on the interaction between the solid Earth and the atmosphere after localized and transient events such as earthquakes and volcanic eruptions.

4.16.3.1 Historical Context

An early example of observations can be traced back to more than a century ago, the eruption of the Krakatoa volcano (Indonesia) on 26 August 1883. After the eruption, coupled

air–sea waves were observed in barographs and tide gages worldwide (Harkrider and Press, 1967). The meteor or comet explosion in Tunguska (Central Siberia) on 30 June 1908 also induced both atmospheric and seismic waves. The latter was similar in amplitude to a $M=5$ earthquake (Whipple, 1930), although the true source of those waves was an explosion in the air, at an estimated altitude of 8 km (Ben-Menahem, 1975).

The interest for the study of such interactions rose significantly later, during Cold War periods, because atmospheric gravity waves were emitted by nuclear explosions and detection and characterization of such waves became an active research field (Hines, 1972). This led to three major advances during the 1960s. First, a better and efficient theoretical description of internal waves in the atmosphere (Hines, 1960), from the ground up to ionospheric heights, was developed. Second, ionospheric sounding networks capable of monitoring the ionospheric response to those waves (Davies, 1962) were deployed. And third, the occurrence of several major earthquakes (Chile, 22 May 1960, M9.5; Alaska, 28 March 1964, M9.2) revealed the generation of internal acoustic waves in the atmosphere by global Rayleigh wave propagation (Donn and Posmentier, 1964).

More recently, significant advances were made toward a quantitative interpretation and prediction of such coupled phenomena. In particular, new types of atmospheric observations have been made possible by the development of GPS (Global Positioning System) ionosphere monitoring (Calais and Minster, 1998; Mannucci et al., 1998) and by the deployment of the International Monitoring System (global seismological, hydroacoustic, radionuclide, and infrasound network aimed at ensuring compliance with the Comprehensive Nuclear-Test-Ban Treaty). Increased accessibility of numerical computations has also made it possible to refine theoretical modeling of coupled systems (Artru et al., 2004; LePichon et al., 2003).

4.16.3.2 Theoretical Preliminaries

4.16.3.2.1 Wave propagation in the atmosphere

After localized events, energy is transmitted in the form of atmospheric internal waves from the surface to the ionosphere. Those waves arise from the interactions of compressional and gravitational forces and are divided into two classes, long-period gravity waves and short-period acoustic waves. The basic physics of acoustic gravity waves was formulated by Hines (1960), and his formalism is now widely used. Let us review the main features of this theory.

We consider an isothermal atmosphere, initially in hydrostatic equilibrium, and include forces from inertia, gravity, and pressure gradients. We assume that disturbances can be regarded as adiabatic process because wave propagation is a sufficiently fast process. We do not include the effects from the rotation of the Earth, and therefore, large-scale tidal and planetary waves are not considered here. We focus on much shorter-wavelength waves.

Let us use the notations that ρ is the density, p is the pressure, v is the neutral gas velocity, g is the gravitational acceleration, and C_s is the constant sound speed. We then have the following three basic equations:Conservation of mass:

$$\frac{\partial \rho}{\partial t} + v\cdot\nabla\rho = -\rho\cdot\nabla v \qquad [7a]$$

Conservation of momentum:

$$\frac{\partial v}{\partial t} + v\cdot\nabla v = g - \frac{1}{\rho}\nabla p \qquad [7b]$$

Adiabaticity:

$$\frac{\partial p}{\partial t} + v\cdot\nabla p = C_s{}^2\left(\frac{\partial p}{\partial t} + v\cdot\nabla p\right) \qquad [7c]$$

In the equilibrium state, $v_0=0$ and both p_0 and p_0 are proportional to exp $(-z/H)$, where $H=C_s^2/\gamma|g|$ is the density scale height and γ is the specific heat ratio. Assuming ρ_1, p_1, and v are small perturbations with no dependency on the y-axis, we may solve the linearized equations as harmonic solutions by assuming that ρ_1, p_1, and v are proportional to $\exp[i(\omega t - k_x x - k_z z)]$. The full dispersion relation takes the form

$$\omega^4 - \omega^2 C_s^2\left(k_x^2 + k_z^2\right) + (\gamma-1)g^2k_x^2 + i\gamma g\omega^2 k_z = 0 \qquad [8]$$

This equation means that, in the presence of gravity, no solution exists in which both k_x and k_z are purely real and different from zero. Let us assume that k_x is real and seek a solution that propagates in the x direction as a harmonic wave. There are now two possibilities, either k_z is purely imaginary or

$$k_z = k_z' + i\frac{\gamma g}{2c^2} = k_z' + i\frac{1}{2H} \qquad [9]$$

The first case (k_z pure imaginary) is appropriate for horizontally propagating surface waves, but it permits no variation of phase with height. The second case is appropriate for traveling disturbances in the atmosphere under the influence of gravity. Thus, the second case is pursued here. The dispersion relation can be rewritten as

$$\omega^2 C_s^2 k_z' 2 = \omega^4 - \omega^2\left(C_s^2k_x^2 + \frac{\gamma^2g^2}{4C_s^2}\right) + (\gamma-1)g^2k_x^2 \qquad [10]$$

For a given horizontal wave number, a real solution for $k_z^{'}$ exists only when the right-hand side of Eqn [7], which is a second-order polynomial in ω, is positive. The two roots ω_1 and ω_2 define the following three different cases:

- If $\omega_1<\omega<\omega_2, k_z'$ is purely imaginary. In this case, we can only have waves trapped at the surface propagating only horizontally (Lamb waves).
- If $\omega<\omega_1<(\gamma-1)^{1/2}g/C_s, k_z'$ is purely real. This corresponds to the internal gravity wave domain, governed primarily by buoyancy. $\omega_g=(\gamma-1)^{1/2}g/C_s$ is the Brunt–Väisälä frequency.
- If $\omega>\omega_2>\gamma g/2C_s, k_z'$ is purely real. This corresponds to the internal acoustic wave domain, governed primarily by compression. The high-frequency limit of those waves is usual sound wave (with exponential increase in amplitude). $\omega_a=\gamma g/2C_s$ is called the acoustic cutoff frequency. In the short-wavelength limit, however, the root ω_2 varies as C_sk_x.

Typically, $\omega_a/2\pi=3.3$ mHz and $\omega_g/2\pi=2.9$ mHz in the lower atmosphere. This classification in acoustic, gravity, and Lamb waves is widely used in aeronomy. It can be extended to nonisothermal models when the scale height is large in comparison with the vertical wavelength. An adaptation of

normal-mode theory to such a nonisothermal model, for the coupled solid Earth–ocean–atmosphere system, has been developed, which allows the simulation of wave propagation for the whole system (Lognonné et al., 1998). The main difference with traditional seismic normal-mode simulation is in the treatment of the upper-boundary condition to reflect the absence of free surface.

Attenuation of waves arises from viscous and thermal dissipations with comparable magnitudes. But except for short-period acoustic waves, attenuation can be neglected up to 100 km of altitude. This is because dominant atmospheric waves that are coupled to Rayleigh waves have frequency content of about 10–20 mHz.

Stronger limitations of this model arise at high altitude because of two reasons; first, amplification with altitude implies that linearized equations will cease to be valid. And second, upward-propagating internal waves eventually reach the ionosphere, where the dynamics is strongly constrained by the influence of magnetic field on charged particles. The ionospheric response to gravity waves is still an object of an extensive literature (Yeh and Liu, 1972).

4.16.3.2.2 Frequency–wavelength domains of interest

Seismic, tsunami, and atmospheric waves result from the action of gravity and elastic forces. Significant effects from both types of forces exist for the relevant frequency range (a few tens of millihertz). Figure 6 represents a normal-mode representation of the whole Earth + ocean + atmosphere system (Artru et al., 2001; Lognonné et al., 1998). The range of existence of seismic, tsunami, acoustic, and gravity waves can be clearly identified, and the different areas where they coexist (seismic and acoustic, tsunami, and gravity) mark the potential for coupled waves. Outside those area, the dynamic coupling is limited to interface waves (e.g., Lamb wave in the atmosphere).

Study of atmospheric signals caused by sources in the solid Earth is mostly based on observations at the surface (seismograms, barograms, tides gages, and infrasound arrays) or at high altitudes through the ionospheric response to upward-propagating internal waves. The latter measurement can only concern the acoustic or gravity frequency ranges. In particular, seismic and tsunami waves are much more likely to produce strong atmospheric signals at high altitude than many other natural or artificial sources. This is because, despite the very small size of displacements at the surface, they present a unique combination of frequency and horizontal wavelength range necessary for an efficient coupling with internal waves in the atmosphere. On the other hand, major energy from ocean swell, located in the same frequency range, may induce some infrasonic signal trapped at the base of the atmosphere (Garces et al., 2003), but will not in general induce internal (i.e., upward-propagating) acoustic waves in the atmosphere, because the wavelength is much shorter than for Rayleigh waves (Arendt and Fritts, 2000).

4.16.3.3 Observation Techniques

4.16.3.3.1 Surface observation: Seismometer, microbarograph, and hydrophones

Both seismic and pressure sensors have been used to characterize the coupling at the Earth surface. In particular, recent worldwide deployment of infrasound arrays for the International Monitoring System has enabled us to detect atmospheric pressure fluctuations related to solid Earth activity (Figure 7). The frequency band for such instruments is typically between 0.1 and 10 Hz (Mutschlecner and Whitaker, 2005). The noise

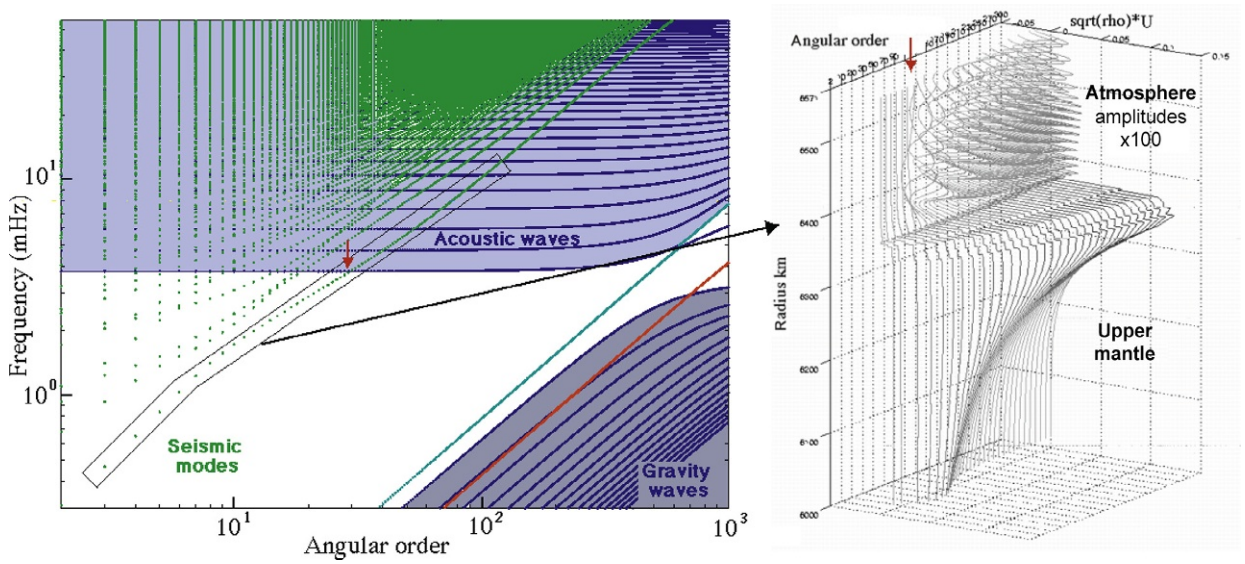


Figure 6 Domains of existence of waves in the solid Earth, ocean, and atmosphere. This figure represents the set of free oscillation modes obtained using standard Minos software (Woodhouse, 1988) for an Earth model composed of the preliminary reference Earth model (PREM) for solid Earth and US Standard Atmosphere 1976 up to 86 km of altitude. Green dots: seismic spheroidal modes; red dots: 'tsunami' mode (corresponding to the 3-km ocean layer in PREM); blue dots: acoustic and gravity mode. As Minos software has a free boundary at the top of the model, some obvious errors exists: the atmospheric mode spectrum is not in general discrete because there is no upper-boundary condition, and furthermore, the lowest acoustic branch shown on this plot would correspond to the nonphysical surface atmospheric mode at the top of the model.

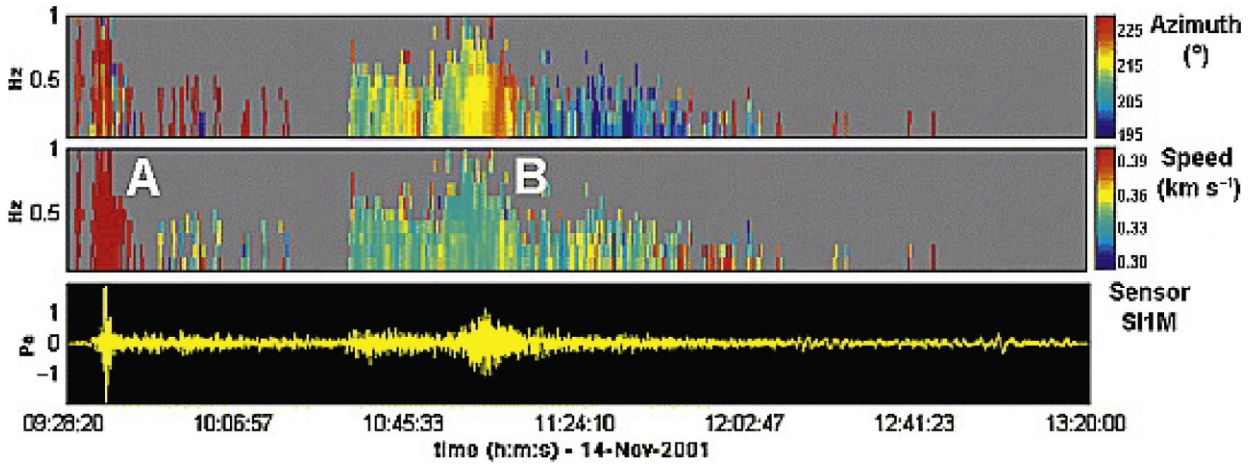


Figure 7 Infrasound record from the I34MN IMS infrasound array (Mongolia) after the 2001 Kunlun earthquake (Northern China, M_s 8.1 on 14 November 2001). Bottom panel: Atmospheric pressure fluctuations filtered between 0.05 and 4 Hz. Top two panels: The color scales indicate the values of the azimuth and the trace velocity measured at the infrasound station. *Y*-axis corresponds to frequencies from 0.05 to 1 Hz. In the first part of the signal, coherent wave trains referred to as Group A are related to local seismically coupled air waves. Due to the coupling at the Earth–air interface, the horizontal trace velocity of the ground-coupled airwaves and the seismic waves is identical (>3 km s^{-1}). Group B corresponds to the propagation of infrasonic waves backscattered by the Kunlun mountain range.

level depends on the wind conditions, from 0.1 to 10 Pa, but it can be reduced by filtering. Those arrays can also be used to determine the azimuth and velocity of the observed signals, which provide information on their origin. Essentially, two types of infrasound signals are observed after earthquakes (LePichon et al., 2002):

- Seismic coupled air waves, essentially a local conversion of seismic wave vertical motion into sound pressure (Donn and Posmentier, 1964).
- Infrasound waves remotely generated. Those can be generated at the epicenter (Bolt, 1964) or backscattered by mountain ranges (Young and Greene, 1982).

In the latter case, infrasounds travel obliquely upward in the atmosphere but are reflected or refracted at various altitudes because of sound velocity variations in the atmosphere (Drob et al., 2003). Such a signal arrives later than the coupled air wave and has different azimuth and velocity. Mutschlecner and Whitaker (2005) showed that amplitude of the observed pressure perturbations depend not only on the epicentral distance and magnitude of the earthquake but also on the stratospheric wind speed, which has a major impact on the propagation of infrasounds.

4.16.3.3.2 High-altitude observation: Ionosondes, Doppler sounding, transmission (GPS), and in situ measurements

Detection of acoustic and gravity waves at intermediate altitudes in the neutral atmosphere is not possible due to the absence of *in situ* measurement and because atmospheric remote sensing generally lacks the resolution that would be needed. However, at higher altitudes, in addition to the exponential amplification of the wave amplitude, the interaction with the local plasma leads to perturbations of the ionosphere that are detectable by using radio sounding techniques.

The ionosphere is the intermediate region between the neutral atmosphere and the magnetosphere, ranging approximately between 60 and 1500 km in altitude. It is a stratified medium, partially ionized by solar radiation (due to ultraviolet light and x-rays). The maximum in electron density is reached between 200 and 400 km of altitude and takes values between 10^5 and 10^7 e^{-1} m^{-3} (**Figure 8**).

Because it is a plasma, the ionosphere has a strong influence on electromagnetic wave propagation. The plasma frequency is the low-frequency cutoff for radio wave propagation and depends on the local electron density N_e as

$$\omega_p = \left(\frac{N_e e^2}{m_e \varepsilon_0}\right)^{\frac{1}{2}} \qquad [11]$$

where e and m_e are the electron charge and mass, respectively, and ε_0 is the permittivity of vacuum. Typical ionospheric plasma frequencies range from 1 to 20 MHz and their maximum occurs at the maximum of ionization. Ionospheric sounding techniques are based on measurement of either the reflection of radio waves below the maximum plasma frequency or the refraction delay of higher-frequency signals transmitted across the ionosphere from satellites.

4.16.3.3.2.1 Reflection

For frequencies lower than 10 MHz, waves are reflected by the ionospheric layer whose plasma frequency (Eqn [11]) exceeds the frequency of signal. One type of radio sounding of the ionosphere is vertical sounding (ionosonde), measuring the travel time of a reflected signal as a function of frequency. This provides information on the electron density profile. Doppler sounding is based on the measurement of the Doppler shift of monochromatic signal, sent vertically upward and reflected back to the ground. The frequency shift between the emitted wave and the reflected wave is directly proportional to the vertical velocity of the reflecting layer. Using such ionospheric sounding network, Blanc (1985) gave a review of natural and artificial sources of signals recorded in those 'ionospheric seismometers.' Both types of sounding can be either ground-based, probing the

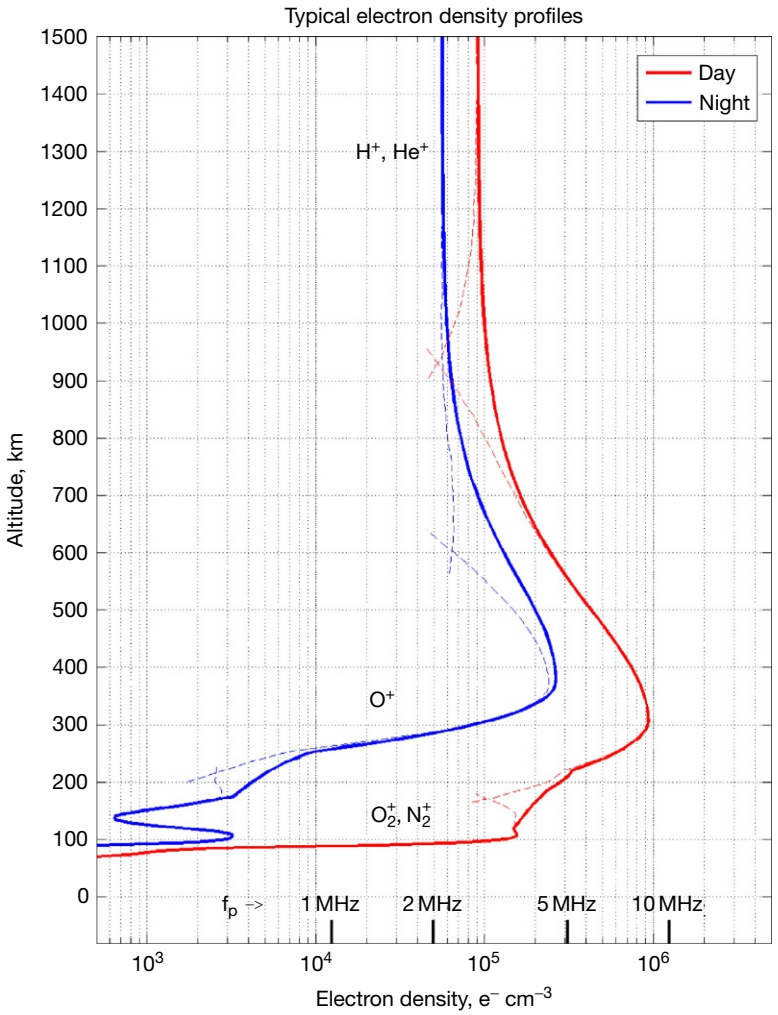


Figure 8 Typical midlatitude electron density profile in the ionosphere. Profiles were obtained using the International Reference Ionosphere model (Bilitza, 2000). Main ions species are indicated (dashed lines are their contribution to ionization), and the plasma frequencies (Eqn [11]) corresponding to different electron densities are indicated.

lower ionosphere up to the electron density maximum, or satellite-based ('top-side sounding').

4.16.3.3.2.2 Transmission

For higher frequencies (typically L-band, 1–2 GHz and higher), electromagnetic waves can propagate across the ionosphere. Ionospheric refraction induces a delay in the travel time, which depends on the frequency of the signal and on the local plasma frequency (Eqn [11]), hence on the local electron density along the ray:

$$\delta_{\rho\text{iono}} = \pm\frac{e^2}{4\pi^2 m_e \epsilon_0}\int_{\text{ray}}\frac{N_e}{f^2}\mathrm{d}l \qquad [12]$$

Using travel time measurements at two different frequencies gives access to the total electron content (TEC), which is the integral of electron density along a path between a radio transmitter and a receiver. This measurement on TEC has become widely accessible from the development in satellite altimetry and positioning, since it is a key correction required to be applied to those systems. In particular, GPS relies on the accurate measurement of satellite–GPS receiver distances and uses a combination of travel time measurements at two frequencies (L1 and L2, respectively, $f_1 = 1.575$ and $f_2 = 1.227$ GHz). Once other sources of error are taken into account (Lanyi and Roth, 1988; Sardon et al., 1994), TEC along the ray can be obtained as a linear combination of the two estimated distances ρ_1 and ρ_2:

$$\text{TEC} = \int_{\text{ray}} N_e\,\mathrm{d}l = \frac{1}{40.3}\frac{f_2^2 - f_1^1}{f_2^2 f_1^2}(\rho_1 - \rho_2) \qquad [13]$$

4.16.3.4 Sources in the Solid Earth

Ionospheric perturbations that follow earthquakes have been observed both near the seismic source and at teleseismic distances (Figure 9). The first published observations were related to the Great Alaskan Earthquake in 1964. Using ionospheric sounding networks, Bolt (1964) and Davies and Baker (1965) observed atmospheric perturbations propagating from the

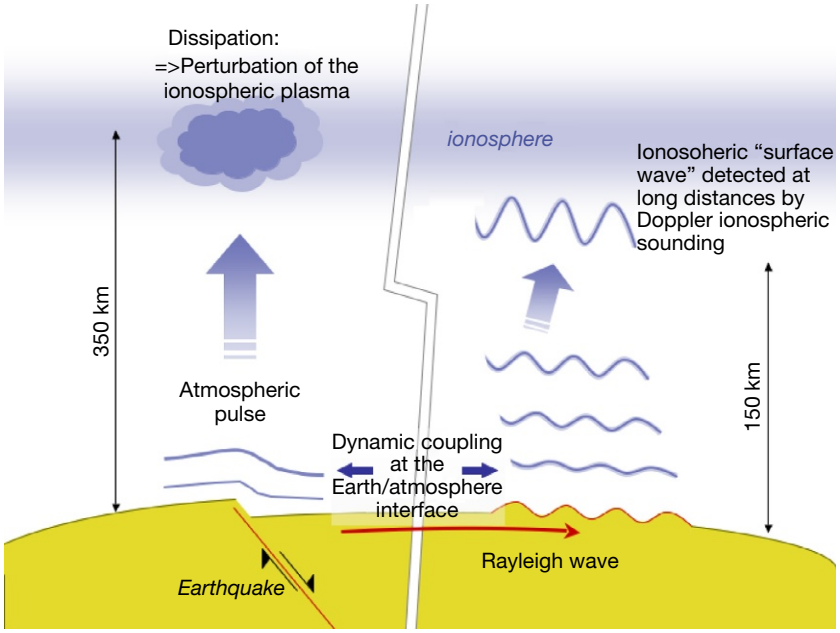


Figure 9 The schematic view of atmospheric waves generated after an earthquake.

epicenter region and the ionospheric signatures of the Rayleigh wave propagation at the sounder location.

4.16.3.4.1 Away from the source: Surface waves and tsunamis

4.16.3.4.1.1 Seismic waves

Seismic surface waves have typical group velocity in the range 3–4 km s^{-1}, whereas the sound velocity in the atmosphere is much smaller (340 m s^{-1}). Therefore, the acoustic wave is sent almost vertically upward and reaches the ionosphere with a delay of 815 min after the passage of Rayleigh waves on the ground. Amplification due to density decrease with altitude can reach 10^4–10^5, and attenuation is only observed for short-period signal above 100 km of altitude. Large-scale vertical oscillations of the ionospheric layers can be easily monitored using the Doppler sounding networks described in **Section 4.16.3.3.2** (Blanc, 1985; Yuen et al., 1969). **Figure 10** presents a recent example of such measurements. Although monitoring networks are still sparse, such ionospheric oscillations are currently observed systematically for most $M \geq 6.5$ earthquakes worldwide.

Using a normal-mode approach, Artru et al. (2001, 2004) and Occhipinti et al. (2010) had modeled oscillations of ionospheric layers induced by surface Rayleigh waves generated by a major earthquake. Their analysis retrieved source parameters from atmospheric events and showed that atmospheric perturbations generated by surface Rayleigh waves may be amplified by a factor of 10 000 to 100 000 upon attaining heights between 150 and 250 km. Later, Rolland et al. (2011a) investigated the dependence of the coupling efficiency between the solid Earth and the atmosphere on atmospheric conditions. Rayleigh wave-induced ionospheric disturbances were first observed using the dense array of GPS (nowadays, GPS is often rephrased as GNSS (Global Navigation Satellite System) to signify satellite systems for positioning in a general sense) as TEC changes in the Western United States by Ducic et al. (2003) and Garcia et al. (2005) after the 2002 Denali earthquake in Alaska.

4.16.3.4.1.2 Tsunami waves

Tsunami waves are expected to induce a similar type of coupling with the atmosphere; despite their small amplitude compared with ocean swells, they can generate atmospheric gravity waves because of their long wavelengths. A possibility of detecting tsunami by monitoring the ionospheric signature of the induced gravity waves was proposed by Peltier and Hines (1976). They discussed theoretical issues on the coupling and concluded that it is feasible.

Geometry of tsunami–gravity wave coupling is very different from the coupling between seismic and acoustic waves. Tsunami wave is nondispersive, and its velocity depends only on gravity g and water depth d, as the velocity is given by $\sqrt{gd}$. From the gravity wave dispersion equation, one can estimate group velocity of the induced gravity wave. Horizontally, it appears to be very close to typical tsunami wave speed, while vertical component is much slower than sound speed (about 50 m s^{-1}). Depending on the period, it would therefore take from one to a few hours for the gravity wave to reach the ionosphere (in contrast to about 10 min for seismic–acoustic waves). The ionospheric perturbation should be behind the tsunami front, with a delay increasing with altitude. Detailed discussion is given in Occhipinti et al. (2008, 2011).

Early papers that proposed using ionospheric measurements to detect tsunami-generating earthquakes (Najita and Yuen, 1979; Najita et al., 1973) focused on perturbations induced by Rayleigh waves preceding a potentially destructive tsunami. Artru et al. (2005) used ionospheric sounding from a very dense GPS network in Japan (GEONET) to detect the perturbations associated with the arrival of tsunami wave, generated by the 23 June 2001 Peru earthquake. The observed arrival time, wavelengths, and orientation were shown to

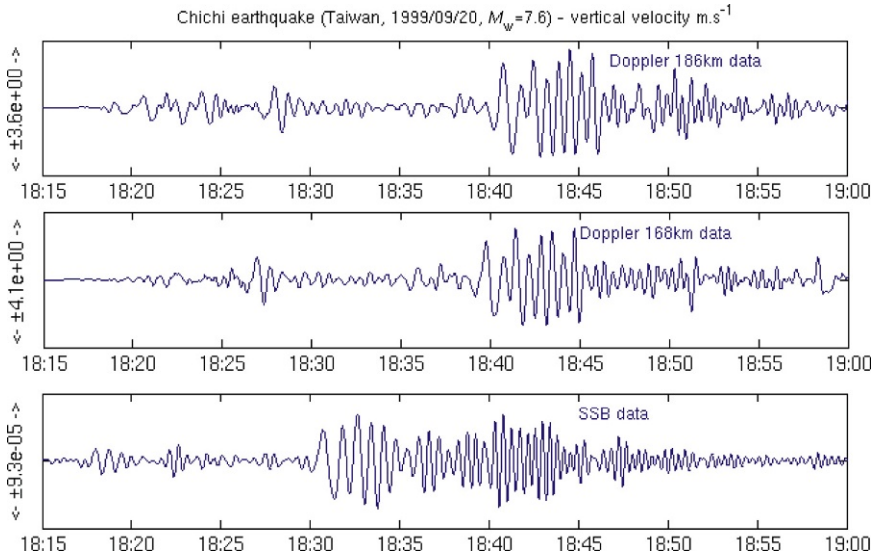


Figure 10 Seismogram and Doppler sounding record (two altitudes: 168 and 186 km) taken in France after Chi–Chi earthquake. Both traces are band-pass filtered between 1 and 50 mHz.

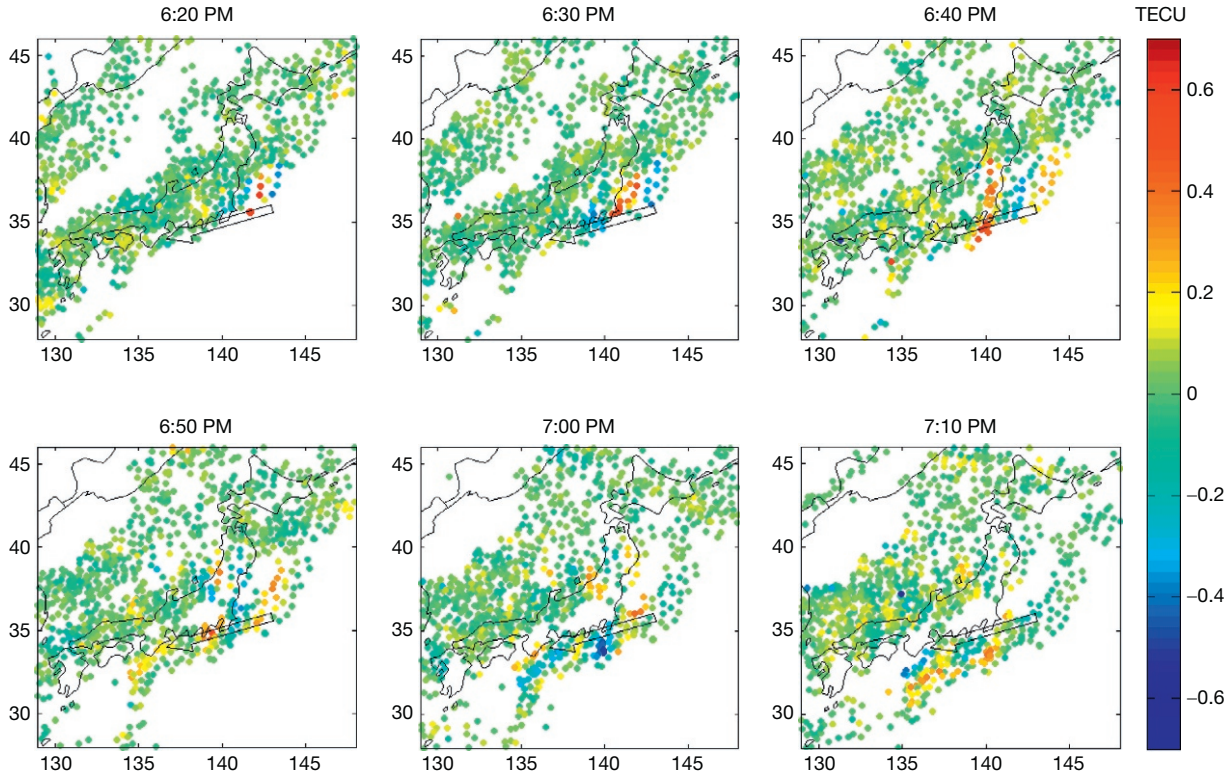


Figure 11 Observed signal for the 23 June 2001 tsunami (initiated offshore Peru): TEC variations plotted at the ionospheric piercing points. A wavelike disturbance is propagating toward the coast of Honshu. This perturbation presents the expected characteristics of tsunami-induced gravity waves and arrives approximately at the same time as the tsunami wave itself.

fit well with theoretical predictions estimated by a simple simulation (Figure 11).

4.16.3.4.2 Direct atmospheric waves

Atmospheric perturbations, observed either at the ground level or in the ionosphere, are often generated away from the measurement location. Propagation of the signal is therefore essentially in the atmosphere. In most cases, the origin of the atmospheric disturbance is the ground motion near an earthquake source (or an underground nuclear explosion). Calais and Minster (1995) detected perturbations in the ionospheric TEC above Southern California after the Northridge earthquake

($M = 6.7$, 17 January 1994) using GPS measurements. Davies and Archambeau (1998) developed a direct modeling of these waves for a simple representation of shallow seismic sources, including high-frequency components of the wave-packet and nonlinear effects. Their results confirmed the seismic origin of the signal, observed after the Northridge earthquake.

The displacement field generated at the Earth's surface produces a piston-like impulse on the atmosphere. Afraimvich et al. (2001) proposed a model for the atmospheric perturbation in the form of 'shock–acoustic waves.' Drob et al. (2003), Il-Young et al. (2002), and Virieux et al. (2004) focused on solving the acoustic wave propagation in order to model data from the CTBT (Comprehensive Nuclear Test Ban Treaty) verification network.

Ionospheric disturbances due to direct acoustic waves were observed by a dense GPS array in Japan after the 2003 Tokachi-Oki earthquake (M_w 8.0) (Heki and Ping, 2005). Their apparent propagation velocity was much slower (up to 1 km s^{-1}) and could be easily distinguished from the disturbances of Rayleigh wave origin. The interaction with geomagnetic field produced directivity of their propagation. Heki and Ping (2005) suggested it was southward (northward component was suppressed by geomagnetic field) in the midlatitude region of the northern hemisphere. For this earthquake, Rolland et al. (2010) later identified Rayleigh waves in addition to the direct acoustic waves. Watada et al. (2006) analyzed the microbarograph records collocated with seismometers for periods 10–50 s and observed pressure disturbances caused by the acoustic coupling between the atmosphere and the ground beneath the sensors.

Using the Japanese GPS array, after the 1994 Hokkaido-Toho-Oki earthquake (M_w 8.3), Astafyeva et al. (2009) clearly identified two components with distinct velocity contrast, that is, the faster Rayleigh wave component and the slower acoustic wave component. This 1994 earthquake occurred within the subducting slab. Also, two large earthquakes occurred in the central Kuril arc that had different focal mechanisms, that is, a shallow-angle thrust earthquake in 2006 and an outer rise normal-fault earthquake in 2007. Astafyeva and Heki (2009) studied the coseismic ionospheric disturbances of these three earthquakes and found that their waveforms (e.g., polarity of the initial changes) depended on their focal mechanisms.

4.16.3.4.2.1 Other sources of infrasounds

In addition to the coupled surface waves and the direct acoustic shock waves, some atmospheric perturbations related to earthquakes can be observed as scattered waves away from the source. For example, a mountain range can become a source of diffracted infrasound waves at the time of passage of Rayleigh waves (LePichon et al., 2002, 2003; Mutschlecner and Whitaker, 2005).

Other 'solid Earth' sources of atmospheric infrasounds include avalanches and rockfall (Bedard, 2000), volcanic activity (Graces, 1997), and chemical explosions and mining blasts (Hagerty et al., 2002).

4.16.3.5 The Case of the Great Sumatra–Andaman Earthquake

The great ($M_w = 9.1$) Sumatra–Andaman earthquake and tsunami of 26 December 2004 (Lay et al., 2005) provided a unique opportunity to study solid Earth–ocean–atmosphere coupling. Indeed, by itself, this particularly large thrust event would generate not only seismic waves and their associated atmospheric Rayleigh waves but also direct atmospheric perturbations. In addition, it triggered an exceptionally large tsunami that propagated across the Indian Ocean. Satellite altimetry observation detected the open-ocean wave, with peak-to-peak amplitude of 40–50 cm (Song et al., 2005). Tide gauge stations also provided tsunami data, allowing Tanioka et al. (2006) to estimate the rupture process somewhat independently from purely seismic data. The propagation of the tsunami wave across the Indian Ocean lasted several hours, which was enough, according to the tsunami–gravity wave coupling theory described earlier, to generate ionospheric signals. Let us review some recent observations published on the subject.

Le Pichon et al. (2005) and Garces et al. (2005) analyzed infrasound array data. They observed distinct packets of signal arriving successively. The first signal was the pressure perturbation generated at the sensor location by the seismic waves, with horizontal trace velocity greater than 3 km s^{-1}. The second signal was an infrasonic wave train with a mean trace velocity of 0.35 m s^{-1} and a dominant period of 10 s, associated with infrasound radiated from the epicenter region. The third signal consisted of large coherent infrasonic wave, similar in velocity to the previous train, with a dominant period of 30 s. Back azimuth reconstruction indicated a source area extending from the northern tip of Sumatra to the northern margin of the Bay of Bengal. This source area suggests that those infrasonic waves were generated by the tsunami wave, either through the interaction of tsunami with the shoreline or as tsunami reached shallow water and generated shorter-wavelength signals (a few tens of kilometers, comparable to 30 s infrasound waves). Mikumo et al. (2008) analyzed the very low-frequency acoustic gravity waves using microbarographs around the Indian Ocean and in Japan and compared with synthetic barograms generated for a realistic thermal structure in the atmosphere up to 220 km.

At higher altitudes, ionospheric perturbations related to the Sumatra earthquake and tsunami have been reported as well. Liu et al. (2006a) used Doppler sounding network in Taiwan, monitoring the vertical motion of a specific ionospheric layer, and detected two distinct disturbances interpreted as the Rayleigh waves, then to the direct acoustic gravity waves emitted by the crustal motion around the earthquake. Other works published were based on continuous GPS data (i.e., the detection of perturbation in the integrated electron content of the ionosphere), although the distribution of permanent receivers in the Indian Ocean region was quite scarce at the time. Heki et al. (2006) studied the signals related to the direct acoustic gravity waves from the source region and used such observations to retrieve information on the rupture process, in particular the rupture propagation speed. Using the same GPS stations in Sumatra and Thailand, Otsuka et al. (2006) studied the variations of the TEC perturbations between different stations and interpreted the variations as the consequence of directivity in the ionospheric response with respect to the neutral atmosphere perturbation. DasGupta et al. (2006) reported smooth variations in TEC detected by GPS stations located on the east coast of India without giving yet any specific interpretation on their origin. Iyemori et al. (2005) reported a rather different type of observation, using ground-based

fluxgate magnetometer. They observed localized, long-period geomagnetic pulsations in Thailand shortly after the origin time of the earthquake and speculated that they were due to the resonant interaction of magnetic field lines with the upward-propagating magnetosonic waves emitted from the earthquake area. Changes in geomagnetism were also observed in space: Balasis and Mandea (2007) detected 30 s pulsation in geomagnetic field associated with the 2004 Sumatra–Andaman and the 2005 Nias earthquakes with a magnetometer on board a low Earth orbiter.

Resonant atmospheric oscillation is considered to occur at frequencies 3.7 and 4.4 mHz (Nishida et al., 2000). Choosakul et al. (2009) found that the atmospheric resonance in these frequencies continued for a few hours in GPS-TEC data taken at Phuket, Thailand, after the 2004 Sumatra–Andaman earthquake.

Searching for the ionospheric disturbances related to the tsunami wave propagation, two studies reported some successful observation. Liu et al. (2006b) used GPS data from five permanent receivers in the southern Indian Ocean and detected traveling ionospheric disturbances with the period in the range 10–20 min, with horizontal propagation speed consistent with the theory of tsunami–gravity wave coupling. Occhipinti et al. (2006) took advantage of the simultaneous sea surface height and TEC measurements provided by altimetry satellites Jason-1 and TOPEX/Poseidon, hence giving data in the open ocean, away from possible coastal perturbation. In addition, they were able to perform a direct three-dimensional (3-D) modeling of the tsunami-generated gravity waves taking into account its interaction with the ionospheric plasma. The simulated and observed perturbations agreed remarkably well considering large uncertainties inherent in combined data sets, models, and theories from very different fields of Earth sciences.

Mai and Kiang (2009) showed that the tsunami-driven gravity wave propagated at the same velocity with the tsunami. Hickey et al. (2009) suggested that the tsunami-driven gravity waves may attain vertical displacements of 2–5 km and electron density perturbations up to 100%. Occhipinti et al. (2008) studied the dependence of the ionospheric signatures of tsunami on geomagnetic field. Hickey et al. (2010) simulated the fluctuations in atmospheric airglow by such gravity waves, which were observed later after the 2011 Tohoku-Oki earthquake (see next chapter).

A suite of large earthquakes have followed the 2004 Sumatra–Andaman earthquake. They include the 2005 M_w 8.6 Nias earthquake (Briggs et al., 2006) and the 2007 M_w 8.5 Bengkulu earthquake (Gusman et al., 2010). Ionospheric disturbances by these earthquakes were studied by Cahyadi and Heki (2013) using a GPS network in Sumatra. The 2007 Bengkulu earthquake showed typical ionospheric responses including the acoustic gravity waves propagating northward with atmospheric resonance at about 5 mHz. However, vigorous plasma bubble activities prohibited near-field observation of ionospheric disturbances by the 2005 Nias earthquake (Cahyadi and Heki, 2012). Later, its ionospheric disturbance was observed in far field by the high-frequency Doppler observations and by an over-the-horizon (OTH) radar system in France, called Nostradamus, by Occhipinti et al. (2010). Coïsson et al. (2011) studied how tsunami-induced gravity wave signature appears in OTH radar data.

Ionospheric disturbances by the 2006 Kuril, 2009 Samoa, and the 2010 Chile earthquake tsunamis were reported by Rolland et al. (2010) and Galvan et al. (2011) using GPS data taken in Hawaii, the Western United States, and Japan. In April 2012, a large earthquake (M_w 8.6) occurred off the coast of northern Sumatra (Meng et al., 2012). Studying its ionospheric disturbances is important because of its unique focal mechanism, that is, one of the largest strike-slip earthquakes.

4.16.3.6 The Cases of the 2011 Megathrust Earthquakes in Tohoku-Oki, Japan

The 11 March 2012 Tohoku-Oki earthquake (M_w 9.0) was the first M9 class earthquake that occurred where dense geodetic and geophysical sensors were available. A dense network composed of 1200 continuous GPS stations, GEONET (GNSS Earth Observation Network), has been in operation since the 1990s by GSI (Geospatial Information Authority of Japan) and provided spectacular observing results concerning the coupling of the solid Earth, the atmosphere, and the ionosphere. Shortly after the earthquake, rapid reports on various aspects of ionospheric disturbances were published. In the near field, the concentric wave fronts created by internal gravity waves were clearly identified with GEONET stations (**Figure 12**) (Rolland et al., 2011a,b; Tsugawa et al., 2011), also with additional GPS stations in the Asia-Pacific region (Tsai et al., 2011). Their onset time was earlier than the expected arrival time of the internal gravity waves that propagated from the surface to the ionosphere. This suggests that the wave propagated first as acoustic waves and excited internal gravity waves in the upper atmosphere. Spatial distribution of the ionospheric disturbances could accurately pinpoint the center of the seafloor uplift. Astafyeva et al. (2011) showed that the first pulse of the disturbance appeared only about 8 min after the earthquake, consistent with the ravel time of the acoustic wave from the surface to the ionosphere. This was well before the tsunami hit the NE Japan coast and indicates that the GPS array is a promising sensor for the early warning system of tsunamis.

By studying the GPS array data from South Korea, Taiwan, and the Japanese GEONET, Chen et al. (2011) found the geomagnetic origin of directivity in the propagation of the acoustic waves in the ionosphere. Liu et al. (2011) identified Rayleigh wave-induced disturbances propagating with a speed of 2.3–3.3 km s^{-1} using GPS receivers in Taiwan. Rolland et al. (2011a,b) and Galvan et al. (2012) analyzed the GPS data in Japan and identified three propagation speeds in the observed ionospheric disturbances and attributed them to gravity waves induced by tsunami (200–300 m s^{-1}) and direct acoustic waves (1000 m s^{-1}) and those excited by the Rayleigh waves (3400 m s^{-1}). Far-field data of the tsunami-induced ionospheric disturbances were, for the first time, observed by an airglow camera located in Hawaii (Makela et al., 2011), and Occhipinti et al. (2011) modeled these signals by internal gravity waves.

The atmospheric resonant oscillations were found to have lasted for 4 h by Saito et al. (2011) and Rolland et al. (2011b). They identified the 3.7 and 5.3 mHz oscillations in addition to the dominant 4.5 mHz oscillations, which were consistent with the frequencies of the atmospheric modes (Kobayashi,

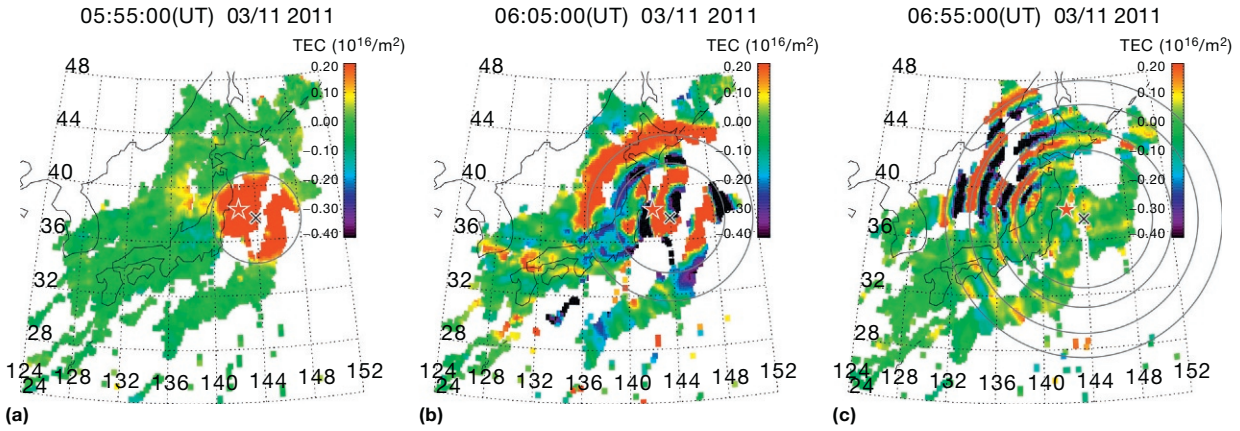


Figure 12 Two-dimensional maps of the detrended TEC at 05:55 UT (a), 06:05 UT (b), and 06:55 UT (c) on 11 March 2011. They show snapshots taken at about 10, 20, and 70 min after the earthquake, respectively. The star and cross marks represent the epicenter and the center of the ionospheric disturbance, respectively. Gray circles represent concentric circles from the center. A movie of the detrended TEC maps with 30 s resolution is available at the NICT website (http://www.seg.nict.go.jp/2011TohokuEarthquake/). Reproduced from Tsugawa T, Saito A, Otsuka Y, et al., (2011) Ionospheric disturbances detected by GPS total electron content observation after the 2011 off-the-Pacific coast of Tohoku Earthquake. *Earth, Planets and Space* 63: 875–879.

2007). These interpretations were substantiated by numerical simulations performed by Matsumura et al. (2011). This study was based on the 2-D simulation method of Shinagawa et al. (2007), which calculated the response of nonlinear, non-hydrostatic, compressible, and neutral atmosphere to impulsive vertical surface motion.

Various non-GPS sensors also detected disturbances in the troposphere and the ionosphere. A microbarograph in Oshu city, Iwate, successfully recorded infrasounds excited in the sea directly above the rupture area (Arai et al., 2011). This is considered to have propagated along the surface–atmosphere boundary as the Lamb wave. The lack of dispersion in the waveforms enabled the preservation of the original offshore waveforms of tsunami. An OTH SuperDARN in Hokkaido, Japan, also detected the ionospheric disturbances caused by this earthquake (Nishitani et al., 2011). Conventional sensors such as ionosonde were also used to study the ionospheric disturbances (Liu and Sun, 2011; Maruyama et al., 2011, 2012).

Several new phenomena were found after this earthquake. Postseismic formation of the ionospheric electron depletion (tsunamigenic ionospheric hole) was found and modeled by Saito et al. (2011) and Kakinami et al. (2012). TEC changes starting 1 h before megathrust earthquakes were suggested by Heki (2011). The same phenomenon was found to have occurred prior to the 2007 Bengkulu earthquake, Indonesia, by Cahyadi and Heki (2013).

4.16.3.7 Sources in the Atmosphere

Atmospheric events that generate infrasounds or gravity waves radiate in some cases enough energy to induce seismic signals through a dynamic coupling, and the inverse energy flow from the atmosphere to the solid Earth can occur. Two main categories are discussed in the succeeding text: volcanic eruptions and sonic booms. The latter include data from the shuttle, supersonic jets, or meteors that enter into the atmosphere.

4.16.3.7.1 Eruptions

Atmospheric waves produced by the Krakatoa eruption in 1883 were observed worldwide by barometric measurements. The propagation of such atmospheric disturbances is similar to those produced by nuclear explosions, and investigations were made in the 1960s to characterize these waves and their impact on the ionosphere (Harkrider and Press, 1967; Row, 1967). Following the eruption of Mount St. Helens in 1980, Roberts et al. (1982) detected a long-lived, large-scale traveling ionospheric disturbance in TEC measurements. Bolt and Tanimoto (1980) reported airwaves in barographic records that circled more than once around the globe.

The eruption of Mount Pinatubo (Philippines) on 15 June 1991 provided a remarkable example of the interaction between the solid and gaseous envelopes of the Earth system. The energy released in several explosions is estimated to be more than 100MT (TNT), generating significant atmospheric pressure waves (see **Figure 6**). Signals related to these waves were observed worldwide in barographs, ionograms, Doppler soundings, TEC measurements, and seismic data, resulting in multitude of analyses for different data. Igarashi et al. (1994) used the Japanese ionospheric observation network to determine the characteristics of the gravity wave and the associated traveling ionospheric disturbances. Kanamori et al. (1994) investigated the source mechanism of atmospheric oscillations from both barographic and seismographic records. Watada and Kanamori (2009) performed theoretical investigation on the resonance between the atmosphere and the lithosphere for periods 230 and 270 s that may occur following volcanic explosions. Johnson (2003) gave a review of infrasound observations emitted from volcanic eruptions and pointed out that it is useful to distinguish regular seismic signals related to subsurface seismicity from the seismicity associated with gas release.

Just like ionospheric disturbances caused by earthquakes, analysis of GPS-TEC data has been shown useful to study ionospheric disturbances by volcanic eruptions. Heki (2006) inferred the explosion energy associated with the September

2004 eruption of Mount Asama using GEONET data, by comparing the amplitudes of the disturbances with those by a mine blast of known explosion energy (Calais et al., 1998). Later, Dautermann et al. (2009a,b) improved the method to estimate the energy and applied it to the 2003 eruption of the Soufrière Hills volcano in the Lesser Antilles.

4.16.3.7.2 Sonic boom

Seismic and underwater perturbations induced by aircraft sonic booms have been studied since the mid-1960s (Cook et al., 1972). The frequency range is much higher than those for previously discussed phenomena. It should be noted that these sonic booms are one of the rare examples of controlled sources. The main effects of sonic booms arise in the waveforms of the strain with the shape of the N-wave overpressure and also in air-coupled Rayleigh wave trains following each N-wave transient. Seismic waves generated by the coupling can then propagate, faster than the original shock wave (Ishihara et al., 2003). More recently, the Concorde (LePichon et al., 2003) and the Space Shuttle provided further opportunities of studying those effects (Sorrells et al., 2002). Using seismic records of the sonic boom by the space shuttle Columbia, returning to the California Edwards Air Force base, Kanamori et al. (1991) observed P-wave pulse across the Los Angeles Basin, probably excited through the motion of high-rise buildings in response to the sonic boom. Yamamoto et al. (2011) observed audible sound, infrasound, and seismic data after the reentry of the Hayabusa sample capsule and the fragmented parts of the spacecraft in 13 June 2010.

Sonic boom was also shown to induce ionospheric perturbations through the atmosphere–ionosphere coupling using GPS data (Calais and Minster, 1998). Water vapor in the rocket exhaust causes large-scale electron depletions due to chemical reaction. Because their signatures are dominant in GPS-TEC data (Furuya and Heki, 2008; Ozeki and Heki, 2010), GPS is not necessarily a desirable tool to study acoustic disturbances, especially by ascending objects.

4.16.4 Conclusion

It has been a traditional practice for seismologists to ignore the atmosphere, the outermost layer of the Earth. To a large extent, this has been justified because of its relatively small effect on seismograms.

On the other hand, because of the existence of the atmospheric layer, we receive some benefits. In addition to understanding the causes of noise, which shed some light on the mechanism in the atmosphere–ocean–solid Earth interactions, one of the recent benefits has been the development of the correlation technique that allows us to retrieve Green's functions among pairs of stations. We can now obtain Green's functions for the solid Earth from the portions of seismograms that do not contain earthquake signals. This owes to the existence of noise and the development of theory on diffuse wavefield generated by noise.

Extending the analysis to include the atmospheric layer will increase opportunities to detect signals that are potentially useful for quantitative analysis of solid Earth processes. Volcanic eruption is an obvious process as it directly generates atmospheric waves. Direct tsunami observation is also possible as surface displacement of ocean is not small and there are atmospheric waves that are coupled to this phenomenon. Shallow earthquakes also emit waves into the atmosphere, which may become a useful source of information. We have probably scratched only the surface of a large body of useful phenomena in the atmosphere, and as we develop the understanding of waves in the whole Earth system, we may discover many things that seismologists have missed in the past 100 years.

References

Afraimvich EL, Perevalova NP, Plotnikov AV, and Uralov AM (2001) The shock-acoustic waves generated by earthquakes. *Annals of Geophysics* 19: 395–409.

Aki K (1957) Space and time spectra of stationary stochastic waves, with special reference to microtremors. *Bulletin of the Earthquake Research Institute* 35: 415–456.

Aki K (1965) A note on the use of microseisms in determining the shallow structures of the Earth's crust. *Geophysics* 30: 665–666.

Arai N, Iwakuni M, Watada S, Imanishi Y, Murayama T, and Nogami M (2011) Atmospheric boundary waves excited by the tsunami generation related to the 2011 great Tohoku-Oki earthquake. *Geophysical Research Letters* 38: L00G18. http://dx.doi.org/10.1029/2011GL049146.

Ardhuin F, Stutzmann E, Schimmel M, and Mangeney A (2011) Ocean wave sources of seismic noise. *Journal of Geophysical Research* 116: C09004. http://dx.doi.org/10.1029/2011JC006952.

Arendt S and Fritts DC (2000) Acoustic radiation by ocean surface waves. *Journal of Fluid Mechanics* 415: 1–21.

Artru J, Ducic V, Kanamori H, Lognonné P, and Murakami M (2005) Ionospheric detection of gravity waves induced by Tsunamis. *Geophysical Journal International* 160(3): 840–848.

Artru J, Farges T, and Lognonné P (2004) Acoustic waves generated from seismic surface waves: Propagation properties determined from Doppler sounding observations and normal-mode modeling. *Geophysical Journal International* 158(3): 1067–1077.

Artru J, Lognonné P, and Blanc E (2001) Normal modes modeling of post-seismic ionospheric oscillations. *Geophysical Research Letters* 28(4): 697–700.

Astafyeva E and Heki K (2009) Dependence of waveform of near-field coseismic ionospheric disturbances on focal mechanisms. *Earth, Planets and Space* 61: 939–943.

Astafyeva E, Heki K, Kiryushkin V, Afraimovich E, and Shalimov S (2009) Two-mode long-distance propagation of coseismic ionosphere disturbances. *Journal of Geophysical Research* 114: A10307. http://dx.doi.org/10.1029/2008JA013853.

Astafyeva E, Lognonné P, and Rolland L (2011) First ionosphere images for the seismic slip of the Tohoku-oki Earthquake. *Geophysical Research Letters* L22104. http://dx.doi.org/10.1029/2011GL049623.

Balasis G and Mandea M (2007) Can electromagnetic disturbances related to the recent great earthquakes be detected by satellite magnetometers? *Tectonophysics* 431: 173–195.

Beauduin R, Lognonné P, Montagner JP, Cacho S, Karczewski JF, and Morand M (1996) The effects of the atmospheric pressure changes on seismic signals or how to improve the quality of a station. *Bulletin of the Seismological Society of America* 86(6): 1760–1769.

Bedard AJ (2000) Atmospheric infrasound. *Physics Today* 53: 32–37.

Ben-Menahem A (1975) Source parameters of the Siberia explosion of June, 1908, from analysis and synthesis of seismic signals at four stations. *Physics of the Earth and Planetary Interiors* 11: 1–35.

Beroza G and Jordan T (1990) Searching for slow and silent earthquakes using free oscillations. *Journal of Geophysical Research* 95: 2485–2510.

Bilitza D (2000) International reference ionosphere 2000. *Radio Science* 36(2): 261–275.

Blanc E (1985) Observations in the upper atmosphere of infrasonic waves from natural or artificial sources: A summary. *Annals of Geophysics* 3(6): 673–688.

Bolt B (1964) Seismic air waves from the great Alaska earthquake. *Nature* 202: 1095–1096.

Bolt BA and Tanimoto T (1980) Atmospheric oscillations after the May 18, 1980 eruption of Mount St. Helens. *Eos* 62(23): 529.

Boy JP, Hinderer J, and Gegout P (1998) Global atmospheric loading and gravity. *Physics of the Earth and Planetary Interiors* 109: 161–177.

Briggs RW, Sieh K, Meltzner AJ, et al. (2006) Deformation and slip along the Sunda Megathrust in the great 2005 Nias-Simeulue earthquake. *Science* 311: 1897–1901.

Bromirski PD (2001) Vibrations from the "Perfect Storm". *Geochemistry, Geophysics, Geosystems* 2: 1030, Paper number: 2000GC000119.

Bromirski PD and Duennebier FK (2002) The near-coastal microseism spectrum: Spatial and temporal wave climate relationships. *Journal of Geophysical Research* 107(B8): 2166. http://dx.doi.org/10.1029/2001JB000265.

Bromirski PD, Flick RE, and Graham N (1999) Ocean wave height determined from inland seismometer data: Implications for investigating wave climate change in the NE Pacific. *Journal of Geophysical Research* 104(C9): 20753–20766.

Bromirski PD and Gerstoft P (2009) Dominant source regions of the Earth's "hum" are coastal. *Geophysical Research Letters* 36: L13303. http://dx.doi.org/10.1029/2009GL038903.

Bromirski PD, Sergienko OV, and MacAyeal DR (2010) Transoceanic infragravity waves impacting Antarctic ice shelves. *Geophysical Research Letters* 37: L02502. http://dx.doi.org/10.1029/2009GL041488.

Bromirski PD and Stephen RA (2012) Response of the Ross Ice Shelf to ocean gravity wave forcing. *Annals of Glaciology* 53: 163–172.

Cahyadi MN and Heki K (2013) Ionospheric disturbances of the 2007 Bengkulu and the 2005 Nias earthquakes, Sumatra, observed with a regional GPS network. *Journal of Geophysical Research* 118: 1–11. http://dx.doi.org/10.1002/jgra.50208.

Calais E and Minster B (1995) GPS detection of ionospheric perturbations following the January 17, 1994, Northridge earthquake. *Geophysical Research Letters* 22(9): 1045–1048.

Calais E and Minster B (1998) GPS, earthquakes, the ionosphere, and the Space Shuttle. *Physics of the Earth and Planetary Interiors* 105(3–4): 167–181.

Calais E, Minster JB, Hofton MA, and Hedlin H (1998) Ionospheric signature of surface mine blasts from Global Positioning System measurements. *Geophysical Journal International* 132: 191202.

Campillo M and Paul A (2003) Long range correlations in the diffuse seismic coda. *Science* 299: 547–549.

Capon J (1972) Long-period signal processing results for LASA, NORSAR and ALPA. *Geophysical Journal of the Royal Astronomical Society* 31: 279–296.

Cessaro RK (1994) Sources of primary and secondary microseisms. *Bulletin of the Seismological Society of America* 84(1): 142–148.

Chen CH, Saito A, Lin CH, et al. (2011) Long-distance propagation of ionospheric disturbance generated by the 2011 off the Pacific coast of Tohoku Earthquake. *Earth, Planets and Space* 63: 881–884.

Choosakul N, Saito A, Iyemori T, and Hashizume M (2009) Excitation of 4-min periodic ionospheric variations following the great Sumatra-Andaman earthquake in 2004. *Journal of Geophysical Research* 114: A10313. http://dx.doi.org/10.1029/2008JA013915.

Cook JC, Gorforth T, and Cook RK (1972) Seismic and underwater responses to sonic boom. *Journal of the Acoustical Society of America* 51(2): 729–741.

Cooper RIB and Longuet-Higgins MS (1951) An experimental study of the pressure variations in standing water waves. *Proceedings of the Royal Society of London A* 206: 424–435.

Coïsson P, Occhipinti G, Lognonné P, and Rolland LM (2011) Tsunami signature in the ionosphere: The innovative role of OTH radar. *Radio Science* 46: RS0D20. http://dx.doi.org/10.1029/2010RS004603.

Crossley DJ, Jensen OG, and Hinderer J (1995) Effective barometric admittance and gravity residuals. *Physic of the Earth and Planetary Interiors* 90: 221–241.

DasGupta A, Das A, Hui D, Bandyopadhyay KK, and Sivaraman MR (2006) Ionospheric perturbations observed by the GPS following the December 26th, 2004 Sumatra-Andaman earthquake. *Earth, Planets and Space* 58(2): 167–172.

Dautermann T, Calais E, Lognonné P, and Mattioli GS (2009a) Lithosphere-atmosphere-ionosphere coupling after the 2003 explosive eruption of the Soufriere Hills Volcano, Montserrat. *Geophysical Journal International* 179: 1537–1546.

Dautermann T, Calais E, and Mattioli GS (2009b) Global Positioning System detection and energy estimation of the ionospheric wave caused by the 13 July 2003 explosion of the Soufriere Hills Volcano, Montserrat. *Journal of Geophysical Research* 114: B02202. http://dx.doi.org/10.1029/2008JB005722.

Davies K (1962) The measurement of ionospheric drifts by means of a Doppler shift technique. *Journal of Geophysical Research* 67: 4909–4913.

Davies JB and Archambeau CB (1998) Modeling of atmospheric and ionospheric disturbances from shallow seismic sources. *Physics of the Earth and Planetary Interiors* 105: 183–199.

Davies K and Baker DM (1965) Ionospheric effects observed around the time of the Alaskan earthquake of March 28, 1964. *Journal of Geophysical Research* 70: 1251–1253.

Dolenc D, Romanowitz B, McGill P, and Wilcock W (2008) Observations of infragravity waves at the ocean-bottom broadband seismic stations Endeavour (KEBB) and Explorer (KXBB). *Geochemistry, Geophysics, Geosystems* 9: Q05007. http://dx.doi.org/10.1029/2008GC001942.

Donn W and Naini B (1973) Sea wave origin of microbaroms and microseisms. *Journal of Geophysical Research* 78(21): 4482–4488.

Donn W and Posmentier ES (1964) Ground-coupled air waves from the great Alaskan earthquake. *Journal of Geophysical Research* 69: 5357.

Drob DP, Picone JM, and Garces M (2003) Global morphology of infrasound propagation. *Journal of Geophysical Research* 108(D21): 4680.

Ducic V, Artru J, and Lognonné P (2003) Ionospheric remote sensing of the Denali Earthquake Rayleigh surface waves. *Geophysical Research Letters* 30: 1951. http://dx.doi.org/10.1029/2003GL017812.

Duvall TL, Jefferies SM, Harvey JW, and Pomerantz M (1993) Time–distance helioseismology. *Nature* 362: 430–432.

Dziewonski AM and Anderson DL (1981) Preliminary reference Earth model. *Physics of the Earth and Planetary Interiors* 25: 297–356.

Ekström G (2001) Time domain analysis of Earth's long-period background seismic radiation. *Journal of Geophysical Research* 106(B11): 26483–26493.

Friedrich A, Kruger F, and Klinge K (1998) Ocean-generated microseismic noise located with the Grafenberg array. *Journal of Seismology* 2: 47–64.

Fukao Y, Nishida K, Suda N, Nawa K, and Kobayashi N (2002) A theory of the earth's background free oscillations. *Journal of Geophysical Research* 107(B9): 2206. http://dx.doi.org/10.1029/2001JB000153.

Furuya T and Heki K (2008) Ionospheric hole behind an ascending rocket observed with a dense GPS array. *Earth, Planets and Space* 60: 235–239.

Galvan DA, Komjathy A, Hickey MP, et al. (2012) Ionospheric signatures of Tohoku-Oki tsunami of March 11, 2011: Model comparisons near the epicenter. *Radio Science* 47: 4003. http://dx.doi.org/10.1029/2012RS005023.

Galvan DA, Komjathy A, Mickey MP, and Mannucci AJ (2011) The 2009 Samoa and 2010 Chile tsunamis as observed in the ionosphere using GPS total electron content. *Journal of Geophysical Research* 116: A06318. http://dx.doi.org/10.1029/2010JA016204.

Garces M, Caron P, Hetzer C, et al. (2005) Deep infrasound from the Sumatra earthquake and tsunami. *Eos, Transactions of the American Geophysical Union* 86(35): 317–320.

Garces M, Hetzer C, Merrifield M, Willis M, and Aucan J (2003) Observations of surf infrasound in Hawaii. *Geophysical Research Letters* 30(24): 2264–2267. http://dx.doi.org/10.1029/2003GL018614.

Garcia R, Crespon F, Ducic V, and Lognonné P (2005) Three-dimensional ionospheric tomography of post-seismic perturbations produced by the Denali earthquake from GPS data. *Geophysical Journal International* 163: 1049–1064.

Gerstoft P and Tanimoto T (2007) A year of microseisms in Southern California. *Geophysical Research Letters* 34: L20304. http://dx.doi.org/10.1029/2007GL031091.

Goldreich P and Keeley AD (1977) Solar Seismology II. The stochastic excitation of the solar P-modes by turbulent convection. *Astrophysical Journal* 212: 243–251.

Graces MA (1997) On the volcanic waveguide. *Journal of Geophysical Research* 102(B10): 22547–22564.

Grevemeyer I, Herber R, and Essen H-H (2000) Microseismological evidence for a changing wave climate in the northeast Atlantic Ocean. *Nature* 408: 349–352.

Gusman AR, Tanioka Y, Kobayashi T, Latief H, and Pandoe W (2010) Slip distribution of the 2007 Bengkulu earthquake inferred from tsunami waveforms and InSAR data. *Journal of Geophysical Research* 115: B12316. http://dx.doi.org/10.1029/2010JB007565.

Gutenberg B (1958) Two types of microseisms. *Journal of Geophysical Research* 63: 595–597.

Hagerty MT, Kim WY, and Martysevich P (2002) Infrasound detection of large mining blasts in Kazakhstan. *Pure and Applied Geophysics* 159: 1063–1079.

Harkrider D and Press F (1967) The Krakatoa air–sea waves: An example of pulse propagation in coupled systems. *Geophysical Journal of the Royal Astronomical Society* 13: 149–159.

Hasselman K (1963) A statistical analysis of the generation of microseisms. *Reviews of Geophysics* 1(2): 177–210.

Haubrich RA, Munk W, and Snodgrass FE (1963) Comparative spectra of microseisms and swell. *Bulletin of the Seismological Society of America* 53: 27–37.

Heki K (2006) Explosion energy of the 2004 eruption of the Asama Volcano, central Japan, inferred from ionospheric disturbances. *Geophysical Research Letters* 33: L14303. http://dx.doi.org/10.1029/2006GL026249.

Heki K (2011) Ionospheric electron enhancement preceding the 2011 Tohoku-Oki earthquake. *Geophysical Research Letters* 38: L17312. http://dx.doi.org/10.1029/2011GL047908.

Heki K, Otsuka Y, Choosakul N, Hemmakorn N, Komolmis T, and Maruyama T (2006) Detection of ruptures of Andaman fault segments in the 2004 great Sumatra earthquake with coseismic ionospheric disturbances. *Journal of Geophysical Research* 111: B09313. http://dx.doi.org/10.1029/2005JB004202.

Heki K and Ping J (2005) Directivity and apparent velocity of the coseismic traveling ionospheric disturbances observed with a dense GPS network. *Earth and Planetary Science Letters* 236: 845–855.

Hickey MP, Schubert G, and Walterscheid RL (2009) Propagation of tsunami-driven gravity waves into the thermosphere and ionosphere. *Journal of Geophysical Research* 114: A08304. http://dx.doi.org/10.1029/2009JA014105.

Hickey MP, Schubert G, and Walterscheid RL (2010) Atmospheric airglow fluctuations due to a tsunami-driven gravity wave disturbance. *Journal of Geophysical Research* 115: A06308. http://dx.doi.org/10.1029/2009JA014977.

Hillers G, Graham N, Campillo M, Kedar S, Landés M, and Shapiro N (2012) Global oceanic microseism sources as seen by seismic arrays and predicted by wave action models. *Geochemistry, Geophysics, Geosystems* 2(13): Q01021. http://dx.doi.org/10.1029/2011GC003875.

Hinderer J and Crossley D (2000) Time variations in gravity and inferences on the Earth's structure and dynamics. *Surveys in Geophysics* 21: 1–45.

Hines CO (1960) Internal atmospheric gravity waves at ionospheric heights. *Canadian Journal of Physics* 38: 1441–1481.

Hines CO (1972) Gravity-waves in atmosphere. *Nature* 239: 73–78.

Igarashi K, Kaminuma S, Nishimuta I, et al. (1994) Ionosphere and atmospheric disturbances around Japan caused by the eruption of Mount-Pinatubo on 15 June, 1991. *Journal of Atmospheric and Terrestrial Physics* 56: 1227–1234.

Il-Young C, Jun M-S, Jeon J-S, and Min KD (2002) Analysis of local seismo-acoustic events in the Korean peninsula. *Geophysical Research Letters* (29). http://dx.doi.org/10.1029/2001GL014060.

Ishihara Y, Tsukada S, Sakai S, Hiramatsu Y, and Furumoto M (2003) The 1998 Miyako fireballs trajectory determined from shock wave records of a dense seismic array. *Earth, Planets and Space* 55: 9–12.

Iyemori T, Nose M, Han D, et al. (2005) Geomagnetic pulsations caused by the Sumatra earthquake on December 26, 2004. *Geophysical Research Letters* 32: L20807. http://dx.doi.org/10.1029/2005GL024083.

Johnson JB (2003) Generation and propagation of infrasonic airwaves from volcanic explosions. *Journal of Volcanology and Geothermal Research* 121: 1–14.

Kakinami Y, Kamogawa M, Tanioka Y, et al. (2012) Tsunamigenic ionospheric hole. *Geophysical Research Letters* 39: L00G27. http://dx.doi.org/10.1029/2011GL050159.

Kanamori H, Mori J, Anderson DL, and Heaton TH (1991) Seismic excitation by the space-shuttle Columbia. *Nature* 349: 781–782.

Kanamori H, Mori J, and Harkrider D (1994) Excitation of atmospheric oscillations by volcanic eruptions. *Journal of Geophysical Research* 22: 21947–21961.

Kedar S, Longuet-Higgins MS, Webb FH, Graham N, Clayton RW, and Jones CE (2008) The origin of deep ocean microseisms in the North Atlantic Ocean. *Proceedings of the Royal Society A* 464(2091): 777–793. http://dx.doi.org/10.1098/rspa.2007.0277.

Kobayashi N (2007) A new method to calculate normal modes. *Geophysical Journal International* 168: 315–331.

Kobayashi N and Nishida K (1998) Continuous excitation of planetary free oscillations by atmospheric disturbances. *Nature* 395: 357–360.

Lacoss RT, Kelly EJ, and Toksoz MN (1969) Estimation of seismic noise structure using seismic arrays. *Geophysics* 34(1): 21–38.

Landau LD and Lifshitz EM (1987) *Fluid Mechanics*. Elmford, New York, USA: Pergamon Press.

Lanyi GE and Roth E (1988) A comparison of mapped and measured total ionospheric electron content using global positioning system and beacon satellite observation. *Radio Science* 23: 483–492.

Lay T, Kanamori H, Ammon CJ, et al. (2005) The great Sumatra–Andaman earthquake of 26 December 2004. *Science* 308(5725): 1127–1133. http://dx.doi.org/10.1126/science.1112250.

Le Pichon A, Herry P, Mialle P, et al. (2005) Infrasound associated with 2004–2005 large Sumatra earthquakes and tsunami. *Geophysical Research Letters* 32: L19802. http://dx.doi.org/10.1029/2005GL023893.

Lee AW (1935) On the direction of approach of microseismic waves. *Proceedings of the Royal Society of London. Series A: Mathematical and Physical Sciences* 149(866): 183–199.

LePichon A, Guilbert J, Vallee M, Dessa JX, and Ulziibat M (2003) Infrasonic imaging of the Kunlun mountains for the great 2001 China earthquake. *Geophysical Research Letters* 30: 1814.

LePichon A, Guilbert J, Vega A, Garces M, and Brachet N (2002) Ground-coupled air waves and diffracted infrasound from the Arequipa earthquake of June 23, 2001. *Geophysical Research Letters* 29: 1886–1889.

Liu JY, Chen CH, Lin CH, Tsai HF, Chen CH, and Kamogawa M (2011) Ionospheric disturbances triggered by the 11 March 2011 M9.0 Tohoku earthquake. *Journal of Geophysical Research* 116: A06319. http://dx.doi.org/10.1029/2011JA016761.

Liu JY and Sun YY (2011) Seismo-traveling ionospheric disturbances of ionograms observed during the 2011 Mw 9.0 Tohoku Earthquake. *Earth, Planets and Space* 63: 897–902.

Liu JY, Tsai YB, Chen SW, et al. (2006a) Giant ionospheric disturbances excited by the M9.3 Sumatra earthquake of 26 December 2004. *Geophysical Research Letters* 33: L02103. http://dx.doi.org/10.1029/2005GL023963.

Liu J-Y, Tsai Y-B, Ma K-F, et al. (2006b) Ionospheric GPS total electron content (TEC) disturbances triggered by the 26 December 2004 Indian Ocean tsunami. *Journal of Geophysical Research* 111: A05303. http://dx.doi.org/10.1029/2005JA011200.

Lobkis O and Weaver R (2001) On the emergence of the Green's function in the correlations of a diffuse field. *Journal of the Acoustical Society of America* 110: 3011–3017.

Lognonné P, Clevede E, and Kanamori H (1998) Computation of seismograms and atmospheric oscillations by normal-mode summation for a spherical Earth model with realistic atmosphere. *Geophysical Journal International* 135: 388–406.

Longuet-Higgins MS (1950) A theory of the origin of microseisms. *Philosophical Transactions of the Royal Society of London, Series A* 243: 1–35.

Mai CL and Kiang JF (2009) Modeling of ionospheric perturbation by 2004 Sumatra tsunami. *Radio Science* 44: RS3011. http://dx.doi.org/10.1029/2008RS004060.

Makela JJ, Lognonné P, Hebert H, et al. (2011) Imaging and modeling the ionospheric airglow response over Hawaii to the tsunami generated by the Tohoku earthquake of 11 March 2011. *Geophysical Research Letters* 38: L00G02. http://dx.doi.org/10.1029/2011GL047860.

Mannucci AJ, Wilson BD, Yuan DN, Ho CH, Lindqwister UJ, and Runge TF (1998) A global mapping technique for GPS-derived ionospheric electron content measurements. *Radio Science* 33: 565–582.

Maruyama T, Tsugawa T, Kato H, Ishii M, and Nishioka M (2012) Rayleigh wave signature in ionograms induced by strong earthquakes. *Journal of Geophysical Research* 117: A08306. http://dx.doi.org/10.1029/2012JA017952.

Maruyama T, Tsugawa T, Kato H, Saito A, Otsuka Y, and Nishioka M (2011) Ionospheric multiple stratifications and irregularities induced by the 2011 off the Pacific coast of Tohoku Earthquake. *Earth, Planets and Space* 63: 869–873.

Matsumura M, Saito A, Iyemori T, et al. (2011) Numerical simulations of atmospheric waves excited by the 2011 off the Pacific coast of Tohoku Earthquake. *Earth, Planets and Space* 63: 885–889.

Meng L, Ampuero JP, Stock J, Duputel Z, Luo Y, and Tsai VC (2012) Earthquake in a maze: Compressional rupture branching during the 2012 Mw 8.6 Sumatra Earthquake. *Science* 337: 724–726.

Merriam JB (1992) An ephemeris for gravity tide predictions at the nanogal level. *Geophysical Journal International* 108: 415–422.

Mikumo T, Shibutani T, Le Pichon A, et al. (2008) Low-frequency acoustic-gravity waves from coseismic vertical deformation associated with the 2004 Sumatra–Andaman earthquake (Mw = 9.2). *Journal of Geophysical Research* 113: B12402.

Müller T and Zürn W (1983) Observation of gravity changes during the passage of cold fronts. *Journal of Geophysics* 53: 155–162.

Mutschlecner JP and Whitaker RW (2005) Infrasound from earthquakes. *Journal of Geophysical Research* 110(D1): D01108.

Najita K, Weaver PF, and Yuen PC (1973) A tsunami warning system using an ionospheric technique. *Proceedings of the IEEE* 62: 563–567.

Najita K and Yuen PC (1979) Long-period oceanic Rayleigh wave group velocity dispersion curve from the Doppler sounding of the ionosphere. *Journal of Geophysical Research* 84: 1253–1260.

Nawa K, Suda N, Fukao Y, Sato T, Aoyama Y, and Shibuya K (2000a) Incessant excitation of the Earth's free oscillations. *Earth, Planets and Space* 50: 3–8.

Nawa K, Suda N, Fukao Y, et al. (2000b) Incessant excitation of the Earth's free oscillations: Global comparison of superconducting gravimeter records. *Physics of the Earth and Planetary Interiors* 120: 289–297.

Nawa K, Suda N, Sato T, Aoyama Y, and Shibuya K (1998) Incessant excitation of the Earth's free oscillations. *Earth, Planets and Space* 50: 3–8.

Niebauer TM (1998) Correcting gravity measurements for the effects of local air pressure. *Journal of Geophysical Research* 93(No. B7): 7989–7991.

Nishida K, Fukao Y, Watada S, et al. (2005a) Array observation of background atmospheric waves in the seismic band from 1 mHz to 0.5 Hz. *Geophysical Journal International* 162: 824–840.

Nishida K, Kobayashi N, and Fukao Y (2000) Resonant oscillations between the solid earth and the atmosphere. *Science* 287: 2244–2246.

Nishida K, Kobayashi N, and Fukao Y (2002) Origin of Earth's ground noise from 2 to 20 mHz. *Geophysical Research Letters* 29(10). http://dx.doi.org/10.1029/2001/GL013862.

Nishida K, Kobayashi N, and Fukao Y (2005b) Resonant oscillation between the solid Earth and the atmosphere. *Science* 287: 2244–2246.

Nishida K, Montagner J, and Kawakatsu H (2009) Global surface wave tomography using seismic hum. *Science* 326: 112. http://dx.doi.org/10.1026/science.1176389.

Nishitani N, Ogawa T, Otsuka Y, Hosokawa K, and Hori T (2011) Propagation of large amplitude ionospheric disturbances with velocity dispersion observed by the SuperDARN Hokkaido radar after the 2011 on the Pacific coast of Tohoku Earthquake. *Earth, Planets and Space* 63: 891–896.

Occhipinti G, Coïsson P, Makela JJ, et al. (2011) Three-dimensional numerical modeling of tsunami-related internal gravity waves in the Hawaiian atmosphere. *Earth, Planets and Space* 63: 847–851.

Occhipinti G, Dorey P, Farges T, and Lognonné P (2010) Nostradamus: The radar that wanted to be a seismometer. *Geophysical Research Letters* 37: L18104. http://dx.doi.org/10.1029/2010GL044009.

Occhipinti G, Kherani EA, and Lognonné P (2008) Geomagnetic dependence of ionospheric disturbances induced by tsunamigenic internal gravity waves. *Geophysical Journal International* 173: 753–765.

Occhipinti G, Lognonné P, Kherani EA, and Hebert H (2006) Three-dimensional waveform modeling of ionospheric signature induced by the 2004 Sumatra tsunami. *Geophysical Research Letters* 33: L20104. http://dx.doi.org/10.1029/2006GL026865.

Okada H (2003) *The Microtremor survey method*. *Geophysical Monograph Series*, vol. 12. Society of Exploration Geophysicists. Washington D.C.: American Geophysical Union Monograph.

Okihiro M, Guza R, and Seymour R (1992) Bound infragravity waves. *Journal of Geophysical Research* 97(C7): 11453–11469.

Otsuka Y, Kotake N, Tsugawa T, et al. (2006) GPS detection of total electron content variations over Indonesia and Thailand following the 26 December 2004 earthquake. *Earth, Planets and Space* 58(2): 159–165.

Ozeki M and Heki K (2010) Ionospheric holes made by ballistic missiles from North Korea detected with a Japanese dense GPS array. *Journal of Geophysical Research* 115: A09314. http://dx.doi.org/10.1029/2010JA015531.

Peltier WR and Hines CO (1976) Possible detection of tsunamis by a monitoring of ionosphere. *Journal of Geophysical Research* 81: 1995–2000.

Peterson J (1993) Observation and modeling of background seismic noise. *U.S. Geological Survey Open-File Report*. Albuquerque, NM.

Phillips OM (1977) *The Dynamics of the Upper Ocean*. Cambridge: Cambridge University Press.

Posmentier ES (1967) A theory of microbaroms. *Geophysical Journal of the Royal Astronomical Society* 13: 487–501.

Prieto GA, Lawrence JF, and Beroza GC (2009) Anelastic Earth structure from the coherency of the ambient seismic field. *Journal of Geophysical Research* 114: B07303. http://dx.doi.org/10.1029/2008JB006067.

Rabbel W and Zschau J (1985) Static deformations and gravity changes at the Earth's surface due to atmospheric loading. *Journal of Geophysics* 56: 81–99.

Rhie J and Romanowicz B (2004) Excitation of Earth's continuous free oscillations by atmosphere-ocean-seafloor coupling. *Nature* 431: 552–556.

Rickett J and Claerbout J (1999) Acoustic daylight imaging via spectral factorization: Helioseismology and reservoir monitoring. *The Leading Edge* 18: 957–960.

Roberts DH, Klobuchar JA, Fougere PF, and Hendrickson DH (1982) A large amplitude travelling ionospheric disturbance produced by the May 18, 1980, explosion of Mount St. Helens. *Journal of Geophysical Research* 87: 6291–6301.

Rolland LM, Lognonné P, Astafyeva E, et al. (2011a) The resonant response of the ionosphere imaged after the 2011 off-the-Pacific coast of Tohoku Earthquake. *Earth, Planets and Space* 63: 853–857.

Rolland LM, Lognonné P, and Munekane H (2011b) Detection and modeling of Rayleigh waves induced patterns in the ionosphere. *Journal of Geophysical Research* 116: A05320. http://dx.doi.org/10.1029/2010JA016060.

Rolland L, Occhipinti G, Lognonné P, and Loevenbruck A (2010) The 29 September 2009 Samoan tsunami in the ionosphere detected off-shore Hawaii. *Geophysical Research Letters* 37: L17191. http://dx.doi.org/10.1029/2010GL044479.

Roult G and Crawford W (2000) Analysis of 'background' free oscillations and how to improve resolution by subtracting the atmospheric pressure signal. *Physics of the Earth and Planetary Interiors* 121(3–4): 325–338.

Roux P, Sabra KG, Gerstoft P, and Kupperman WA (2005a) P-waves from cross correlation of seismic noise. *Geophysical Research Letters* 32: L19303. http://dx.doi.org/10.1029/2005GL023803.

Roux P, Sabra KG, Gerstoft P, Kupperman WA, and Roux A (2005b) Ambient noise cross correlation in free space: Theoretical approach. *Journal of the Acoustical Society of America* 117: 79–84. http://dx.doi.org/10.1121/1.1830673.

Row RV (1967) Acoustic-gravity waves in the upper atmosphere due to a nuclear detonation and an earthquake. *Journal of Geophysical Research* 72: 1599–1610.

Sabra KG, Gerstoft P, Roux P, and Kupperman WA (2005a) Extracting time-domain Green's function estimates from ambient seismic noise. *Geophysical Research Letters* 32: L03310. http://dx.doi.org/10.1029/2004GL021862.

Sabra KG, Gerstoft P, Roux P, Kupperman WA, and Fehler M (2005b) Surface wave tomography from microseisms in Southern California. *Geophysical Research Letters* 32: L14311. http://dx.doi.org/10.1029/2005GL023155.

Saito T (2010) Love-wave excitation due to the interaction between a propagating ocean wave and the sea-bottom topography. *Geophysical Journal International* 182: 1515–1523. http://dx.doi.org/10.1111/j.1365-246X.2010.04695.x.

Saito A, Tsugawa T, Otsuka Y, et al. (2011) Acoustic resonance and plasma depletion detected by GPS total electron content observation after the 2011 off-the-Pacific coast of Tohoku Earthquake. *Earth, Planets and Space* 63: 863–867.

Santo T (1960) Investigations into microseisms using the observational data of many stations in Japan (Part I) – On the origin of microseisms. *Bulletin of the Earthquake Research Institute* 37: 307–325.

Sardon E, Rius A, and Zarraoa N (1994) Estimation of the transmitter and receiver differential biases and the ionospheric total electron-content from Global Positioning System observations. *Radio Science* 29: 577–586.

Schulte-Pelkum V, Earle PS, and Vernon FL (2004) Strong directivity of ocean-generated seismic noise. *Geochemistry, Geophysics, Geosystems* 5: Q03004. http://dx.doi.org/10.1029/2003GC000520.

Shapiro N, Campillo M, Stehley L, and Ritzwoller M (2005) High-resolution surface-wave tomography from ambient seismic noise. *Science* 307: 1615–1618.

Shinagawa H, Iyemori T, Saito S, and Maruyama T (2007) A numerical simulation of ionospheric and atmospheric variations associated with the Sumatra earthquake on December 26, 2004. *Earth, Planets and Space* 59: 1015–1026.

Snieder R (2004) Extracting the Green's function from the correlation of coda waves: A derivation based on stationary phase. *Physical Review E* 69: 046610.

Song YT, Ji C, Fu L-L, et al. (2005) The 26 December 2004 tsunami source estimated from satellite radar altimetry and seismic waves. *Geophysical Research Letters* 32: L20601. http://dx.doi.org/10.1029/2005GL023683.

Sorrells GG (1971) A preliminary investigation into the relationship between long-period noise and local fluctuations in the atmospheric pressure field. *Geophysical Journal of the Royal Astronomical Society* 26: 71–82.

Sorrells G, Bonner J, and Herrin ET (2002) Seismic precursors to space shuttle shock fronts. *Pure and Applied Geophysics* 159: 1153–1181.

Sorrells G, McDonald J, Der Z, and Herrin E (1971) Earth motion caused by local atmospheric pressure changes. *Geophysical Journal of the Royal Astronomical Society* 26: 83–98.

Spratt RS (1982) Modelling the effect of atmospheric pressure variations on gravity. *Geophysical Journal of the Royal Astronomical Society* 26: 83–98.

Suda N, Nawa K, and Fukao Y (1998) Earth's background free oscillations. *Science* 279: 2089–2091.

Sutton G and Barstow N (1996) Ocean bottom microseisms from a distant supertyphoon. *Geophysical Research Letters* 23(No. 5): 499–502.

Tanimoto T (1999) Excitation of normal modes by atmospheric turbulence: Source of long-period seismic noise. *Geophysical Journal International* 136: 395–402.

Tanimoto T (2005) The oceanic excitation hypothesis for the continuous oscillations of the Earth. *Geophysical Journal International* 160(1). http://dx.doi.org/10.1111/j.1365-246X.2005.02491.x.

Tanimoto T (2007a) Excitation of normal modes by nonlinear interaction of ocean waves. *Geophysical Journal International* 168: 571–582. http://dx.doi.org/10.1111/j.1365-246X.2006.03240.x.

Tanimoto T (2007b) Excitation of microseisms. *Geophysical Research Letters* 34: L05308. http://dx.doi.org/10.1029/2006GL029046.

Tanimoto T and Um J (1999) Cause of continuous oscillations of the Earth. *Journal of Geophysical Research* 104(B12): 28273–28739.

Tanimoto T, Um J, Nishida K, and Kobayashi N (1998) Earth's continuous oscillations observed on seismically quiet days. *Geophysical Research Letters* 25(10): 1553–1556.

Tanioka Y, Yudhicara T, Kususose S, et al. (2006) Rupture process of the 2004 great Sumatra–Andaman earthquake estimated from tsunami waveforms. *Earth, Planets and Space* 58: 203–209.

Tennekes H and Lumley JL (1972) *A First Course in Turbulence*. Cambridge, MA, USA: MIT Press.

Traer J, Gerstoft P, Bromirski P, and Shearer P (2012) Microseisms and hum from ocean surface gravity waves. *Journal of Geophysical Research* 117: B11307. http://dx.doi.org/10.1029/2012JB009550.

Tsai V (2011) Understanding the amplitudes of noise correlation measurements. *Journal of Geophysical Research* 116: B09311. http://dx.doi.org/10.1029/2011JB008483.

Tsai HF, Liu JY, Lin CH, and Chen CH (2011) Tracking the epicenter and the tsunami origin with GPS ionosphere observation. *Earth, Planets and Space* 63: 859–862.

Tsugawa T, Saito A, Otsuka Y, et al. (2011) Ionospheric disturbances detected by GPS total electron content observation after the 2011 off-the-Pacific coast of Tohoku Earthquake. *Earth, Planets and Space* 63: 875–879.

Uchiyama Y and McWilliams JC (2008) Infragravity waves in the deep ocean: Generation, propagation, and seismic hum excitation. *Journal of Geophysical Research* 113: C07029. http://dx.doi.org/10.1029/2007JC004562.

Van Dam TM and Wahr JM (1987) Displacement of the Earth's surface due to atmospheric loading: Effects on gravity and baseline measurements. *Journal of Geophysical Research* 92(B2): 1281–1286.

van Tiggelen BA (2003) Green function retrieval and time reversal in a disordered world. *Physical Review Letters* 91: 243904.

Virieux J, Garnier N, Blanc E, and Dessa JX (2004) Paraxial ray tracing for atmospheric wave propagation. *Geophysical Research Letters* 31: L20106.

Warburton RJ and Goodkind JM (1977) The influence of barometric-pressure variations on gravity. *Geophysical Journal of the Royal Astronomical Society* 48: 281–292.

Watada S and Kanamori H (2009) Acoustic resonant oscillations between the atmosphere and the solid Earth during the 1991 Mt. Pinatubo eruption. *Journal of Geophysical Research* 115: B12319. http://dx.doi.org/10.1029/2010JB007747.

Watada S, Kunugi T, Hirata K, et al. (2006) Atmospheric pressure change associated with the 2003 Tokachi-Oki earthquake. *Geophysical Research Letters* 33: L24306. http://dx.doi.org/10.1029/2006GL027967.

Watada S and Masters G (2001) Oceanic excitation of the continuous oscillations of the Earth, *Eos Transactions. American Geophysical Union* 82(47): SS2A-0620.

Waxler R and Gilbert KE (2006) The radiation of atmospheric microbaroms by ocean waves. *Journal of the Acoustical Society of America* 119(1): 2651–2664. http://dx.doi.org/10.1121/1.2191607.

Weaver R and Lobkis O (2001) Ultrasonics without a source: Thermal fluctuation correlations at MHz frequencies. *Physical Review Letters* 87: 134301.

Webb S, Zhang X, and Crawford W (1991) Infragravity waves in the deep ocean. *Journal of Geophysical Research* 96(C2): 2723–2736.

Whipple FJW (1930) The great Siberian meteor and the waves, seismic and aerial, which it produced. *Quarterly Journal of the Royal Meteorological Society* 56: 287–304.

Widmer-Schnidrig R (2003) What can superconducting gravimeters contribute to normal-mode seismology? *Bulletin of the Seismological Society of America* 93(3): 1370–1380.

Wielandt E and Steim JM (1986) A digital very-broadband seismograph. *Annals of Geophysics* 4B: 227–232.

Woodhouse JH (1988) The calculation of the eigenfrequencies and eigenfunctions of the free oscillations of the Earth and the Sun. In: Doornbos DJ (ed.) *Seismological Algorithms: Computational Methods and Computer Programs*, pp. 321–370. London: Academic Press.

Yamamoto M, Ishihara Y, Hiramatsu Y, et al. (2011) Detection of acoustic/infrasonic/seismic waves generated by hypersonic re-entry of the HAYABUSA capsule and fragmented parts of the spacecraft. *Proceedings of the Astronomical Society of Japan* 63: 971–978.

Yeh KC and Liu CH (1972) *Theory of Ionospheric Waves*. New York: Academic Press.

Young JM and Greene GE (1982) Anomalous infrasound generated by the Alaskan earthquake of 28 March 1964. *Journal of the Acoustical Society of America* 71: 334–339.

Yuen PC, Weaver PF, Suzuki RK, and Furumoto AS (1969) Continuous travelling coupling between seismic waves and the ionosphere evident in May 1968 Japan earthquake data. *Journal of Geophysical Research* 74: 2256–2264.

Zhang J, Gerstoft P, and Bromirski P (2010) Pelagic and coastal sources of P-wave microseisms: Generation under tropical cyclones. *Geophysical Research Letters* 37: L15301. http://dx.doi.org/10.1029/2010GL044288.

Zürn W and Widmer R (1995) On noise reduction in vertical seismic records below 2 mHz using local barometric pressure. *Geophysical Research Letters* 22(24): 3537–5340.

This page intentionally left blank

4.17 Episodic Aseismic Slip at Plate Boundaries

SY Schwartz, University of California Santa Cruz, Santa Cruz, CA, USA

© 2015 Elsevier B.V. All rights reserved.

4.17.1	**Seismic and Aseismic Slip at Plate Boundaries**	445
4.17.2	**Modes of EAS**	447
4.17.2.1	Afterslip	447
4.17.2.2	Preslip	448
4.17.2.3	Interseismic Slip Events	449
4.17.2.4	Slip Episodes	451
4.17.3	**EAS and Seismic Tremor**	451
4.17.4	**Global Observations of EAS**	454
4.17.4.1	Convergent Margins	454
4.17.4.1.1	Cascadia	454
4.17.4.1.2	Southwest Japan	454
4.17.4.1.3	Mexico	455
4.17.4.1.4	New Zealand	456
4.17.4.1.5	Costa Rica	456
4.17.4.1.6	Alaska	456
4.17.4.1.7	Northeast Japan	457
4.17.4.2	San Andreas Fault	457
4.17.4.3	Hawaii	458
4.17.4.4	Oceanic Transform Faults	458
4.17.5	**EAS and Seismicity**	459
4.17.6	**Mechanics of EAS**	460
4.17.7	**Recurrence of EAS**	461
4.17.8	**Conclusions**	461
Acknowledgments		462
References		462

4.17.1 Seismic and Aseismic Slip at Plate Boundaries

Most plate boundaries accommodate far-field motions on localized zones of deformation through aseismic creep at depth and earthquake failure in the upper crust. Earthquake slip and the transition with depth from seismic to aseismic deformation can be understood in terms of the frictional response of materials. Realistic faults have a complicated dependence of friction on slip velocity, time, and slip distance known as rate- and state-dependent friction. If frictional resistance to fault movement decreases faster than stress dissipation because of movement, sliding occurs in sudden slips with associated stress drops (earthquakes), punctuated by periods of no motion as stress recharges. This motion is commonly referred to as stick-slip, and the frictional behavior necessary for it as velocity weakening. If this condition is not met and fault strength does not decrease with slip, stable sliding occurs. The frictional condition for stable sliding is known as velocity strengthening. As temperature increases with depth, a transition from velocity-weakening to velocity-strengthening frictional behavior is believed to be responsible for the termination of interplate earthquakes and accommodation of plate motion through stable sliding or creep. The region of the plate boundary exhibiting velocity-weakening behavior and capable of generating earthquakes is often referred to as the seismogenic zone (**Figure 1**). The seismogenic zone may possess a second frictional transition to velocity strengthening (stable sliding) at its shallow up-dip edge and although earthquakes do not nucleate outside of the seismogenic zone, seismic rupture propagation has been observed to extend into both the up- and down-dip stable sliding frictional regimes. The frictional property of these regions of the fault that can propagate but not nucleate seismic rupture is known as conditional stability.

On time scales of years, strain accumulates at plate boundary seismogenic zones as stress from plate motions builds. When the stress reaches a failure threshold, strain release occurs via large earthquakes. However, several continental strike-slip faults, such as the San Andreas Fault, exhibit segments that release strain fairly continuously or creep both at the surface (Galehouse, 2002) and at seismogenic depths (Schmidt et al., 2005; Wesson, 1988). Creeping segments were first recognized by offsets of cultural features and an abundance of microseismicity in the absence of sizable earthquakes. Although the microseismicity may occur at a high rate, it does not contribute significantly to total fault slip and the majority of the strain in these fault segments is released through aseismic motions. Early instrumental measurements of aseismic motion using conventional geodetic techniques displayed primarily linear slip with time (e.g., Galehouse and Lienkaemper, 2003). Indeed, determinations of aseismic deformation rates are rare, in general, because of the high

 http://dx.doi.org/10.1016/B978-0-444-53802-4.00084-1

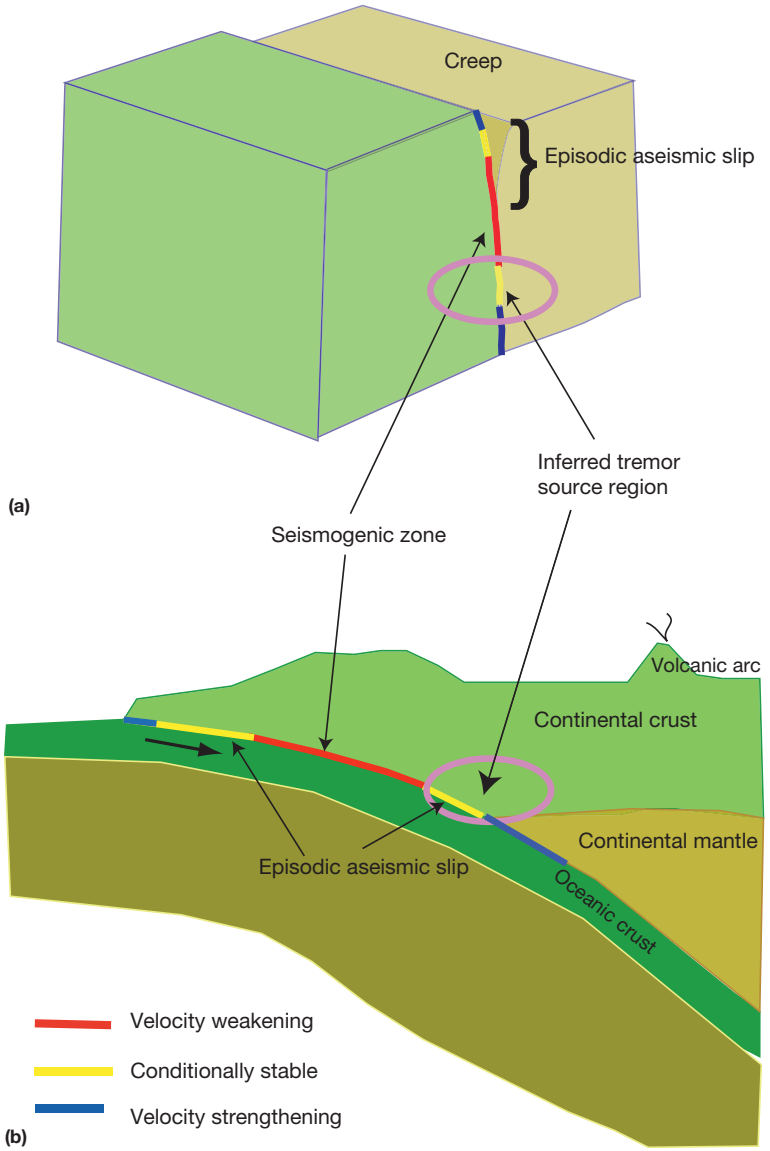


Figure 1 Cartoons of the seismogenic zone at (a) a vertical strike-slip fault and (b) a convergent plate boundary, with frictional properties and regions hosting episodic aseismic slip (EAS) labeled. Earthquakes nucleate in velocity-weakening materials, are capable of propagating into conditionally stable but not velocity-strengthening regions. EAS associated with strike-slip earthquakes is confined to shallow depths but occurs at both the down-dip and up-dip edges of the seismogenic zone at convergent margins. Seismic tremor appears to locate at comparable depths (below the seismogenic zone) in both environments.

detection threshold required to measure such slow processes, resulting in the widely held belief that aseismic motions occur at steady linear rates. Fortunately, the recent acquisition and integration of continuous global positioning system (CGPS), interferometric synthetic aperture radar (InSAR), borehole strainmeter, and creepmeter data at plate boundaries has improved the precision and long-term stability of deformation measurements, revealing significant fluctuations in aseismic deformation rates. These new data have allowed the distinction of several types of aseismic slip processes at plate boundaries including the following: (1) steady-state creep; (2) episodic slip events; (3) afterslip; (4) triggered slip; and (5) precursory slip, each having important implications for the earthquake process. Aseismic episodic slip events are recognized as periods of accelerated slip lasting between a few hours to a few years and have been documented to occur at convergent and strike-slip margins and in continental faulting environments. Afterslip occurs as a rapid acceleration of aseismic slip following a mainshock and is thought to be induced by unrelieved stresses within velocity-strengthening or conditionally stable regions of the fault (Marone et al., 1991). The concentration of afterslip at shallow depths and its logarithmic decay with time distinguish it from postseismic, viscoelastic relaxation where motion, or flow, is typically distributed deeper in the lower crust and upper mantle and continues for years to tens of years, decaying exponentially. Like afterslip, triggered aseismic slip also involves acceleration in slip rate following an earthquake, but it is produced by either dynamic stresses generated during the

passage of large-amplitude surface waves or static stresses from the occurrence of nearby earthquakes on adjacent faults. Precursory slip, although controversial, has been suggested based on observations of accelerated slip to last hundreds of seconds to many years, prior to several earthquakes in a variety of tectonic settings.

Like aseismic motion, earthquakes display variations in their slip rate, having an average value of meters per second, as much as 10 orders of magnitude faster than plate motions and 4–9 orders of magnitude faster than episodic aseismic slip (EAS) rates. **Figure 2** shows a plot of characteristic slip distance versus slip duration, modified from Beroza and Jordan (1990) for seismic and aseismic deformational processes. It shows a continuum of slip velocities with both normal and slow earthquakes radiating seismic energy at the fast end, and steady-state creep and plate motions at the slow end. The existence of slow earthquakes was postulated based on analysis of anomalous free oscillations of Earth (Beroza and Jordan, 1990), anomalous excitation of long-period surface waves (Cifuentes and Silver, 1989; Kanamori and Cipar, 1974), and weak seismic radiation at high frequencies or anomalous low-frequency radiation (Kanamori and Stewart, 1976; Newman and Okal, 1998; Perez-Campos et al., 2003) or strong tsunami relative to seismic excitation (Pelayo and Wiens, 1992). Some examples of slow earthquakes include a predecessor to the great 1960 Chilean earthquake (Kanamori and Cipar, 1974; Linde and Silver, 1989), a successor to the 1978 Izu-Oshima earthquake in Japan (Sacks et al., 1981), and the 1992 Nicaragua tsunami earthquake (Kanamori and Kikuchi, 1993). The relative slowness of these events has been attributed to unusual material properties along the fault interface (e.g., Bilek et al., 2004) such that a large amount of energy is dissipated during deformation, making less of it available for seismic radiation.

EAS refers to deformational processes with rates intermediate between earthquakes and plate motions (**Figure 2**). This mode of strain release was first recognized from dilatational strain records in Japan (Sacks et al., 1971) and has since been documented to occur in a variety of forms in almost all tectonic environments. These observations are allowing a more complete understanding of strain accumulation and release at plate boundaries, with the rate of fault loading and unloading varying in both space and time. Seismic slip at most plate boundaries accounts for only a fraction of plate tectonic motions, so the inclusion of EAS can provide a better quantification of the seismic moment budget. Earthquake slip on one fault segment changes the stress on adjacent segments and may accelerate or decelerate seismic activity. This abrupt stress change may also initiate other modes of strain release, such as EAS. Conversely, EAS also perturbs the surrounding stress field and may either increase or relieve stress on an adjacent fault segment, bringing it closer to or farther from earthquake failure, respectively. Therefore, an improved understanding of EAS and its ability to trigger earthquakes is important for seismic hazard assessment.

This chapter uses distinct nomenclature to distinguish different manifestations of EAS based on when it occurs in the earthquake cycle (i.e., preslip, interseismic slip, and afterslip), and whether it is better characterized as an intermediate-term change in deformation rate (episodic) or a shorter-term, earthquake-like occurrence (event).

4.17.2 Modes of EAS

4.17.2.1 Afterslip

Aseismic slip following large earthquakes, or afterslip, was first documented for the 1966 Parkfield earthquake along the San Andreas Fault (Smith and Wyss, 1968). Since then, it has been detected at many other strike-slip (Bucknam et al., 1978; Williams and Magistrale, 1989) and subduction zone thrust faults. Characteristics of afterslip include (1) location on the coseismic fault plane; (2) slip duration of days to months, with a rapid initial rate of slip followed by a logarithmic decay; and (3) cumulative values that are a significant fraction of the average coseismic slip. Marone et al. (1991) proposed a model for earthquake afterslip in which fault kinematics are prescribed by rate- and state-dependent friction laws. Unrelieved stresses on the fault induce aseismic slip after the

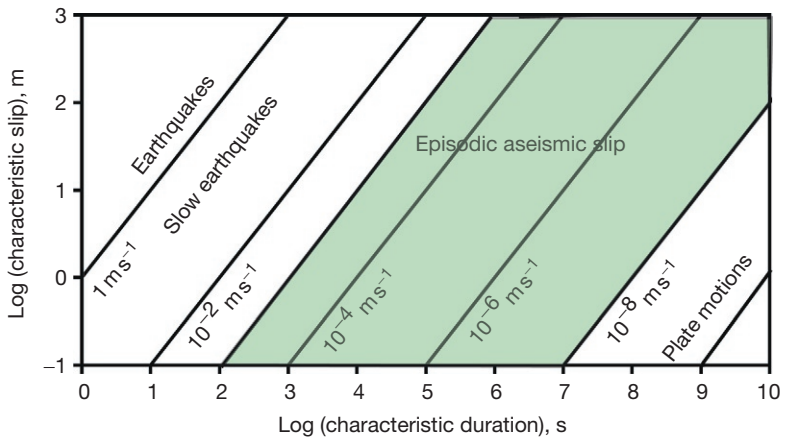


Figure 2 Plot of characteristic slip magnitude versus slip duration for various deformational processes. Both normal and slow earthquakes radiate seismic energy and occur at the fast end while plate tectonic motions occur at the slow end of the velocity spectrum. Most EAS (shaded region) occurs at intermediate velocities. Modified from Beroza GC and Jordan T (1990) Searching for slow and silent earthquakes. *Journal of Geophysical Research* 95: 2485–1510.

earthquake within velocity-strengthening (or conditionally stable) regions of the fault. Their particular model consisted of a thin layer with velocity-strengthening properties, such as fault gouge, overlying a much thicker, velocity-weakening layer. Coseismic rupture initiates and propagates in the lower layer, and is arrested on encountering the shallow velocity-strengthening layer, resulting in a stress concentration at the frictional transition that drives afterslip in the shallow layer. The existence of a slip deficit at shallow depth is required to drive the afterslip. Therefore, in this particular model, if coseismic rupture propagates to the surface, relaxing stress in the velocity-strengthening layer, no afterslip is predicted. In fact, such a phenomenon was observed for the 1992 Landers, California, earthquake where significant coseismic slip extended to the surface (Wald and Heaton, 1994) and negligible afterslip was detected (Sylvester, 1993). The model of Marone et al. (1991) is very well suited to vertical strike-slip faults that develop thick velocity-strengthening gouge layers as they mature. It has successfully reproduced the afterslip time histories for the 1966 Parkfield and 1987 Superstition Hills, California, earthquakes and yielded estimates of the frictional transition depths and other important rheological properties of these fault zones from the logarithmic decay of the afterslip (Marone et al., 1991).

With the increase in GPS monitoring of plate boundaries, numerous observations of afterslip have been made at shallow subduction-zone thrust faults (e.g., Pritchard and Simons, 2006). In this tectonic environment, afterslip is most commonly reported at the downdip edge of the seismogenic zone (Bürgmann et al., 2001; Heki et al., 1997; Hutton et al., 2001; Melbourne et al., 2002; Yagi et al., 2003). This is likely due to the location of most GPS stations on land, directly above the deep plate interface. These network configurations provide little resolution of postseismic deformation occurring offshore. An exception to this is GPS observations following the M_w 8.7 2005 Nias, Indonesia, earthquake, which were made close enough to the trench to allow resolution of shallow afterslip. Eleven months of afterslip, equivalent to an M_w 8.2 earthquake, occurred both up- and down-dip of the mainshock asperity (Hsu et al., 2006; Kreemer et al., 2006). Although substantial afterslip was detected following the M_w 9.2 2004 Sumatra–Andaman earthquake, the lack of near-field deformation data made determination of specific locations of afterslip difficult. The best assessments place afterslip at shallow depth, occurring close to, or possibly overlapping, regions of coseismic slip, with larger displacements occurring in the northern section of the rupture where coseismic slip was relatively small (Subarya et al., 2006; Vigny et al., 2006). Afterslip following the M_w 8.0 2003 Tokachi-oki earthquake occurred within the same depth range as the coseismic slip (within the seismogenic zone), but at unique positions along strike, adjacent to, but not overlapping, the mainshock asperities (Matsubara et al., 2005). Additional examples of great subduction zone earthquakes with afterslip occurring in regions of low coseismic slip include the M_w 8.0 2007 Pisco, Peru (Perfettini et al., 2010), the M_w 8.8 2010 Maule, Chile (Vigny et al., 2011), and the M_w 9.0 2011 Tohoku, Japan (Ozawa et al., 2012), events. In other words, in most cases, afterslip appears to fill in regions of slip deficit left by mainshocks as predicted by state- and rate-dependent frictional models (Marone et al., 1991).

Afterslip is important to document and understand because it often accounts for as much slip as occurs in the mainshock it follows. This large amount of afterslip may help explain the observation that seismic moment at many plate boundaries is far less than the total moment expected if all plate motion is accommodated seismically. Studies of afterslip are also important to test rate- and state-dependent friction laws. Accurate estimation of the decay time of afterslip, as well as rigorous numerical tests of whether the observed decay is, in fact, logarithmic, and remains so over the entire observation period, will allow rate- and state-friction laws to be rigorously tested with field observations.

4.17.2.2 Preslip

Numerical and laboratory models of fault failure exhibit accelerating deformation prior to dynamic slip instabilities (e.g., Dieterich, 1979). However, observations of aseismic deformation rate changes prior to earthquakes are quite rare. Roeloffs (2006) examined available data on aseismic deformation-rate changes prior to earthquakes and reported ten credible accounts of accelerated slip, lasting a few minutes to more than 10 years, preceding earthquakes ranging in magnitude between 3.5 and 9.2. Roeloffs (2006) also documented the absence of such anomalies for seven well-instrumented events with magnitudes between 6.0 and 8.4 and estimated the maximum equivalent preslip moment that may have been released and gone undetected by existing geodetic instrumentation. Anomalous, slow motion immediately before great subduction zone thrust earthquakes was suggested for the M_w 7.9 1944 Tonankai and M_w 8.0 1946 Nankaido, Japan (Linde and Sacks, 2002 and references therein), the M_w 9.5 1960 Chilean (Cifuentes and Silver, 1989; Kanamori and Cipar, 1974), the M_w 9.2 1964 Alaskan (Hamilton and Shennan, 2005), and the M 9 1700 Cascadia (Shennan et al., 1998) events. For these great earthquakes, aseismic preslip was suggested in order to account for the slow deformation detected on leveling, tide gauge, and water well-level data and from microfossil analysis, but details of the slow slip time histories were not determined. The relevant deformation data for the 1944 and 1946 events in southwest Japan were satisfied with 2 m of slow slip on the plate interface downdip of the coseismic slip. Preslip at the down-dip extension of the seismogenic zone was also proposed for the smaller M_w 7.7 1983 Japan Sea thrust earthquake, based on a borehole strainmeter recording of ~100 aseismic strain transient events (Linde et al., 1988). More recently, precursory slip before the largest aftershock of the M_w 8.4 2001 Peru earthquake was detected on a continuous GPS station in Arequipa, Peru, and shown to be consistent with slow slip along the plate interface (Melbourne and Webb, 2002). Two sequences of migrating foreshocks of the great M_w 9 2011 Tohoku-Oki earthquake were interpreted as slow-slip transients propagating toward and leading up to rupture initiation (Kato et al., 2012). Stress loading from the second sequence, which involved large slip rates, was identified as the probable trigger of unstable dynamic rupture of the mainshock. Precursory slow slip that involved 2–3 cm of vertical deformation over 5–7 days was confirmed a few weeks prior to the 2011 Tohoku mainshock by measurements of Absolute Pressure Gauges (APGs) deployed on the seafloor (Ito et al., 2013). The M_w 7.4 March 2012 underthrusting earthquake in southern Mexico was preceded by a geodetically detected slow slip event (SSE) that

migrated along strike to the epicentral location of the mainshock (Graham et al., 2012). Since there are very few near-field seismic and geodetic observations of large subduction zone events, it is still not clear whether these positive examples of precursory slow slip prior to megathrust rupture reflect a common preparation process.

The absence of preseismic deformation rate changes prior to the M_w 6.9 1989 Loma Prieta, the M_w 7.3 1992 Landers, the M_w 7.1 1999 Hector Mines, and the M_w 6.0 2004 Parkfield earthquakes in CA, and the M_w 7.6 1999 Chi Chi, Taiwan, earthquake, which were recorded by borehole strainmeters, dilatometers, and CGPS at various distances from the epicenter, constrains possible preslip moment to be less than equivalent magnitude 5.4, 4.8, 5, 3.2, and 6 earthquakes, respectively (Roeloffs, 2006). The paucity of positive preslip observations may be due, in part, to the existence of signals below the present detection thresholds. The expansion of continuous GPS networks, borehole strainmeters, and InSAR observations at plate boundaries should greatly improve the opportunity to detect and study preslip.

4.17.2.3 Interseismic Slip Events

The interseismic pattern of surface deformation in the vicinity of active faults is generally attributed to strain accumulation on a locked shallow portion of the fault overlying a deeper continuously slipping region. Therefore, geodetic data collected between earthquakes can be used to infer the locking depth of faults as well as the pattern of strain accumulation on the locked portions (e.g., Bürgmann et al., 2000; Norabuena et al., 2004). Locked strike-slip faults produce a rather sharp decrease in horizontal surface velocity, which reaches zero as the fault is approached (Figure 3(a), top). At convergent margins, where the plate interface is locked, the down-going slab drags the overlying plate with it in the direction of plate motion, producing a horizontal velocity field above the coupled plate boundary that parallels the convergence direction and decreases with distance from the trench (Figure 3(b), top). When the strain reaches the failure strength of the fault, an earthquake occurs, and sudden slip on the plate interface causes surface motions that are in the same (Figure 3(a), bottom) and opposite directions (Figure 3(b), bottom) of strain accumulation at strike-slip and subduction zone margins, respectively. Earthquake motion is complete within seconds to minutes, depending on the size of the earthquake. Eventually, the strain accumulation cycle begins again and surface displacements return to their previous directions and rates (Figure 3, bottom).

This strain accumulation and release pattern associated with the earthquake cycle is clearly visible in the daily position

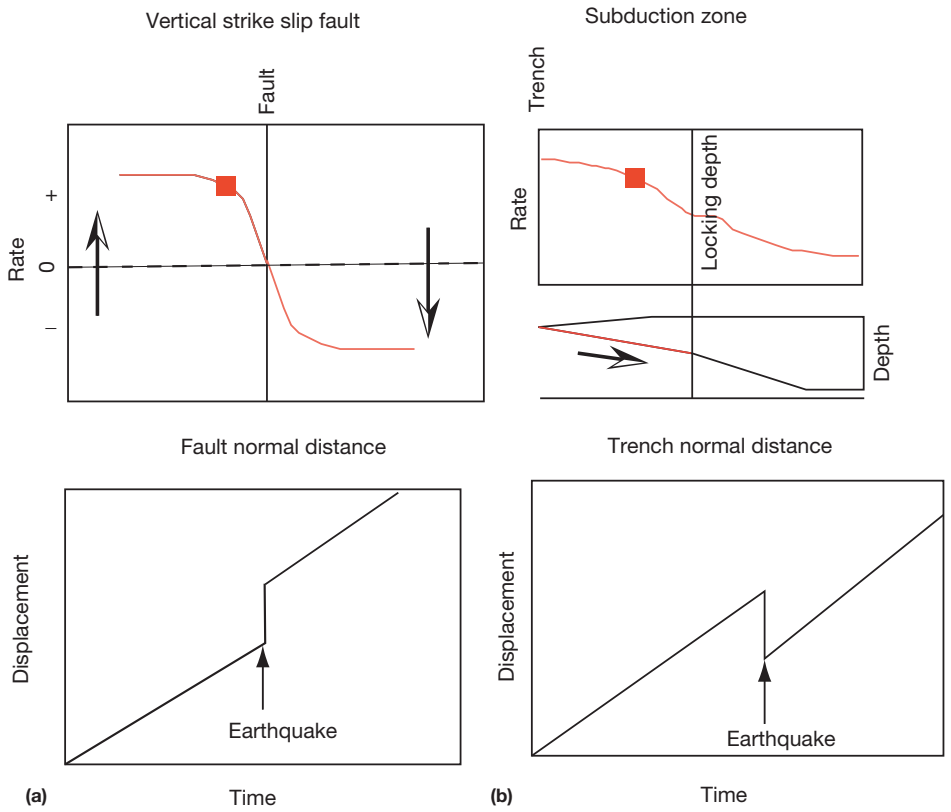


Figure 3 Top row: Schematic geodetic strain accumulation patterns across (a) a vertical strike-slip fault and (b) a fully coupled subduction zone interface. The position of the fault and the horizontal projection of the depth of the transition from a locked to a slipping fault is marked as the solid black lines. The red square corresponds to the location of a hypothetical GPS station for which displacement as a function of time is indicated in the bottom row. Bottom row: Horizontal displacement at the station located with the red square in row 1, showing strain accumulation and coseismic displacement associated with an earthquake, indicated by the arrow.

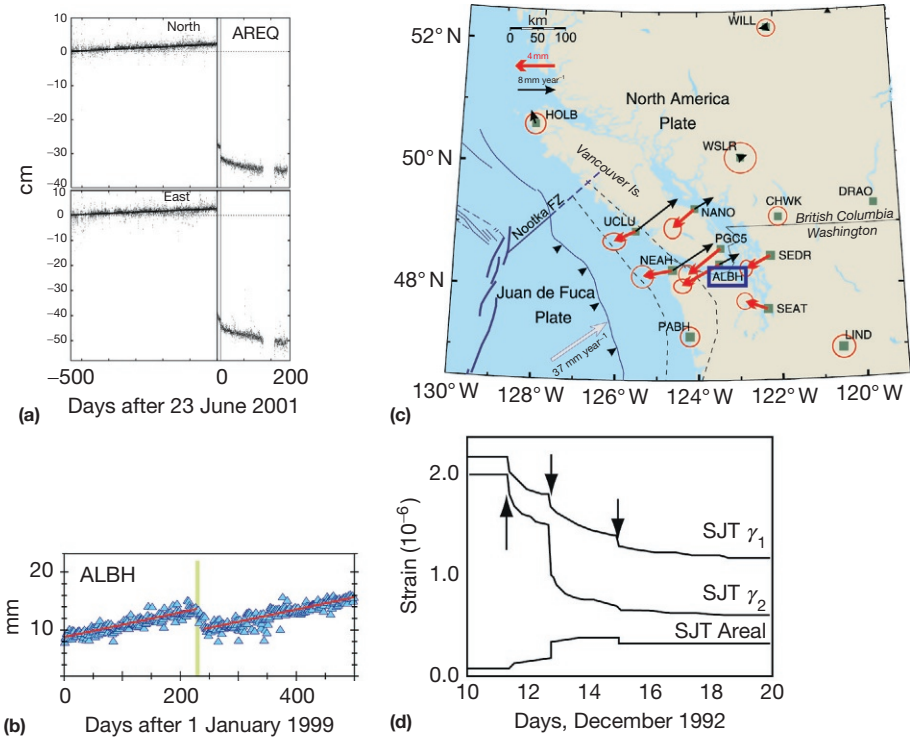


Figure 4 (a) Horizontal position solutions showing interseismic (slow linear northeast motion), coseismic (large, nearly instantaneous southwest motion), and postseismic deformation (slow logarithmic southeast motion) recorded at GPS station Arequipa (AREQ) before and after the M_w 8.4 23 June 2001 Peru earthquake. (b) Longitudinal position solutions at GPS station ALBH, above the northern Cascadia subduction zone, showing a slow slip event (SSE) (marked by the green line) that occurred in August 1999. (c) Map showing the location of the GPS stations that recorded the 1999 SSE. Red arrows show displacements caused by the SSE and black arrows indicate the interseismic strain accumulation velocities. (d) Strain data from strainmeter SJT, located along the San Andreas Fault near San Juan Bautista, during 10 days in December 1992. The arrows indicate the occurrence of SSEs on the fault parallel (γ_1), fault normal (γ_2), and areal strain components. (a) Adapted from Melbourne TI and Webb FH (2002) Precursory transient slip during the 2001 $M_w = 8.4$ Peru earthquake sequence from continuous GPS. *Geophysical Research Letters* 29(21): 2032, http://dx.doi.org/10.1029/2002GL015533. (c) Adapted from Dragert H, Wang K, and James TS (2001) A silent slip event on the deeper Cascadia subduction interface. *Science* 292: 1525–1528. (d) Adapted from Linde AT, Gladwin M, Johnston M, Gwyther R, and Bilham R (1996) A slow earthquake sequence on the San Andreas Fault. *Nature* 383: 65–68.

changes recorded at GPS station Arequipa (AREQ) in Peru above the South American subduction zone (**Figure 4(a)**). The linear northeast motion visible in the 500 days preceding the 2001 M_w 8.4 Peru earthquake is parallel to convergence between the Nazca and South American plates and represents interseismic strain accumulation on the plate interface. On 23 June 2001 (Day 0 in **Figure 4(a)**), a large displacement of 30–40 cm with the opposite sense of motion (southwest) indicates the sudden occurrence of slip on the plate interface, that is, an earthquake. This event was followed by afterslip for several months before the strain accumulation pattern resumed at the same rate as prior to the earthquake. **Figure 4(b)** shows a similar pattern in daily longitudinal position changes at GPS station ALBH on Vancouver Island, overlying the Cascadia subduction zone; however, here the earthquake-like motion that occurs on Day 230 is much smaller (~5 mm) and takes almost 20 days to complete before returning to the strain accumulation signal. This slow, earthquake-like slip that occurs on the plate interface but generates no seismic radiation is often referred to as an SSE. Surface velocities and displacements for the strain accumulation period (black vectors) and this SSE (red vectors), determined at GPS stations operated by the Geological Survey of Canada and part of the Pacific Northwest Geodetic Array, are shown in **Figure 4(c)**. The identification of SSEs in GPS displacement time series that may contain other deformation signals or be contaminated by meteorological effects, local monument motion, and/or reference frame errors, is difficult and requires careful signal processing techniques (e.g., McGuire and Segall, 2003). Borehole strainmeters can provide observations of crustal strain at a much higher sensitivity in the critical SSE time range of months to minutes. **Figure 4(d)** shows strain changes detected in borehole strainmeters near the San Andreas Fault north of San Juan Bautista, CA, during a 10-day period in December 1992. The large coherent strain changes evident between December 11 and 15 have been interpreted as a series of SSEs between 4 and 8 km depth on the fault (Johnston and Linde, 2002; Linde et al., 1996).

Once SSEs have been identified, associated strain or surface displacement data can be modeled, using elastic dislocation theory (Okada, 1985) to determine the SSE location on an assumed fault plane, average slip (d), seismic moment ($M_0 = \mu A d$, in which A is the fault area and μ is the shear modulus of the surrounding material) or equivalent

earthquake moment-magnitude ($M_w = 2/3 \log(M_0) - 6.07$, with M_0 in Nm). In some areas, such as Cascadia, Japan, and the Parkfield region of the San Andreas Fault with dense data coverage, deformation data have been inverted for the non-uniform distribution of slip or slip rate as a function of time on the assumed fault plane. Similar to coseismic slip during earthquakes, interseismic slow slip has been shown to vary over the fault plane with isolated patches of higher or faster slip.

Figure 5 shows the location of all well-documented interseismic SSEs that are discussed in the remainder of this chapter. A large majority of them occur at convergent margins. These plate boundaries generate the largest earthquakes and tend to be relatively well instrumented, so it is not surprising that the majority of SSEs locate here.

4.17.2.4 Slip Episodes

Much of the transient deformation along strike-slip faults in California occurs as deviations in slip rate at depth from long-term rates (slip episodes) (Murray and Segall, 2005; Nadeau and McEvilly, 2004) or as shallow failure of the uppermost several hundred meters of the fault (Gladwin et al., 1994), rather than as SSEs at seismogenic depths. Analysis of strainmeter and creepmeter data for several slip episodes in the Parkfield region of the San Andreas Fault constrained the depth of slip to the upper few hundred meters of the surface (Johnston and Linde, 2002). Similar results have been obtained for other portions of the creeping section of the San Andreas Fault (Evans et al., 1981), which supports the idea that most shallow slip episodes result from soil failure may be triggered by increased rainfall (Roeloffs, 2001), and are, therefore, of little relevance for understanding fault mechanics. However, it has been shown in some cases that shallow slip rates of creeping faults are modulated by slip at depth or slip on adjacent faults. For instance, surface slip accompanied an SSE near San Juan Bautista (Johnston and Linde, 2002; Linde et al., 1996) and changes in surface slip rates were detected on creeping faults following large earthquakes in several locations nearby (Roeloffs, 2001). So deviations in shallow slip rate from long-term rates can be indicators of faulting processes at depth within the seismogenic zone.

4.17.3 EAS and Seismic Tremor

EAS releases accumulated plate boundary strain with durations on the order of minutes to years, too slow to generate seismic radiation. However, Rogers and Dragert (2003) discovered that SSEs in the northern Cascadia subduction zone correlate with episodes of seismic tremor, a signal that is commonly associated with active volcanoes (**Figure 6(a)**). Seismic tremor activity has been associated with SSEs in other subduction zones including southwest Japan (e.g., Hirose and Obara, 2005; Obara, 2010), Alaska (Peterson and Christensen, 2009), Costa Rica (Brown et al., 2005; Outerbridge et al., 2010; Walter et al., 2011), Mexico (Brudzinski et al., 2010; Payero et al., 2008), and New Zealand (Kim et al., 2011). Subduction zone tremor is characterized by long-duration (minutes to hours), low-amplitude seismic signals without clear body wave arrivals. Examples of seismic tremor observed during SSEs in Cascadia, southwest Japan, and southeast Alaska are shown in **Figure 7**. Tremor was originally identified by the coincidence of high-amplitude envelopes on several nearby stations (e.g., McCausland et al., 2005; Obara, 2002). The plots on the right of **Figure 7** show envelope functions for the seismic records shown on the left. These envelopes show that an increased amplitude of tremor can persist for over 30 min. The actual shape of these envelopes is correlated between stations to identify tremor as a relatively widespread phenomenon, rather than as a noise process isolated near a single station. Tremor duration varies between regions and episodes, and is often pulsed in nature (**Figure 7**). Tremor activity has traditionally been measured as the number of hours per day that contain tremors of any duration (**Figure 6**). The actual percentage of each hour filled by tremor varies significantly between different regions and tremor episodes.

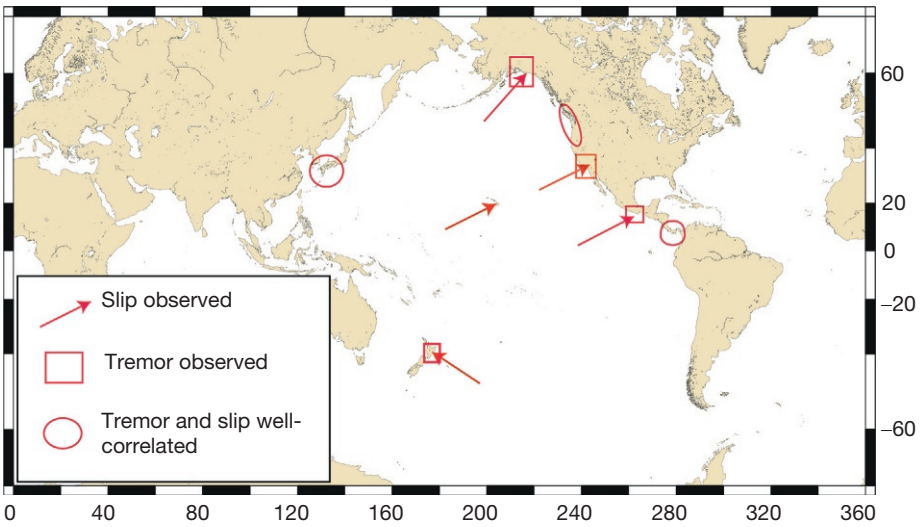


Figure 5 Map showing the location of SSEs discussed in this chapter. Circles and boxes represent regions where seismic tremor has been observed and is or is not well correlated with slow slip, respectively.

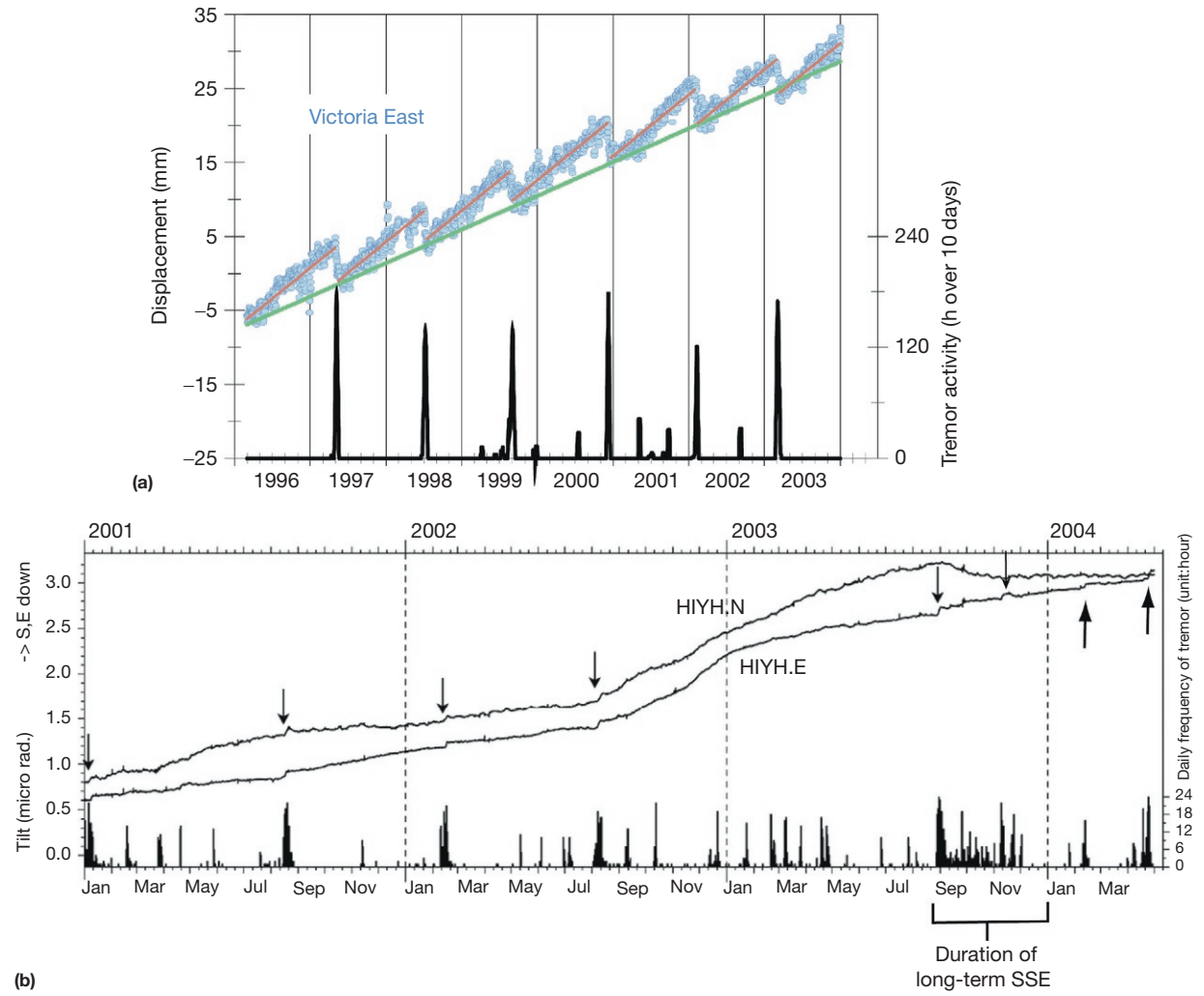


Figure 6 Correlated slow slip and seismic tremor activity in the northern Cascadia (a) and western Shikoku, Japan (b) subduction zones. Bursts in tremor activity, indicated as hours per unit time, strongly correlate with the SSEs apparent on GPS displacement (a) or tiltmeter (b) data. (a) Blue circles show east positions at GPS station ALBH, and the green line indicates the long-term eastward strain accumulation signal, which is interrupted by SSEs (causing motion reversals) about every 13–16 months. From Dragert H, Wang K, and Rogers G (2004) Geodetic and seismic signature of episodic tremor and slip in the northern Cascadia subduction zone. *Earth, Planets and Space* 56: 1143–1150. (b) The north and east tilt records from Hi-net station HIYH, with arrows indicating short-term SSEs. The long-term Bungo channel SSE began in August 2003 and corresponds to the highest concentration of tremor in the studied period. Modified from Obara K and Hirose H (2006) Non-volcanic deep low-frequency tremors accompanying slow slips in the southwest Japan subduction zone. *Tectonophysics* 417(1–2): 33–51.

Like volcanic tremor, which has been associated with the underground movement of fluids, subduction zone tremor is generally characterized as low-frequency, containing peak frequencies that are lower than local earthquake activity. Most tremor shows a rapid dropoff in frequency content above about 5–6 Hz (Kao et al., 2005). Based on the spectral properties of subduction tremor and its presumed similarity to volcanic tremor, early generation models invoked the presence of fluids. Tremor depths in southwest Japan and Cascadia, between 25 and 40 km, coincide with pressure and temperature conditions that are suitable for slab dehydration and tremor sources appear to cluster near regions of high vp/vs ratio consistent with the presence of fluids (Kurashimo and Naoshi, 2004; Matsubara et al., 2005; Shelly et al., 2006). However, Shelly et al. (2006) identified impulsive *P*- and *S*-arrivals embedded in tremor in SW Japan and used them to improve the location of tremor hypocenters. Their improved tremor locations formed a plane parallel to the inferred top of the subducting Philippine plate, more consistent with shear dislocation than fluid movement. By using well-located tremor events as templates, Shelly et al. (2007) systematically searched tremor signals to detect matching waveforms at multiple stations. This technique revealed a nearly continuous sequence of detected events during times of active tremor, establishing that non-volcanic tremor beneath Shikoku, Japan can be explained by a swarm of low frequency earthquakes (LFEs), with each event most likely caused by shear failure on the plate interface. The quality discrepancy between Japanese borehole stations and surface stations used in North America delayed the progress in identifying individual tremor sources in Cascadia.

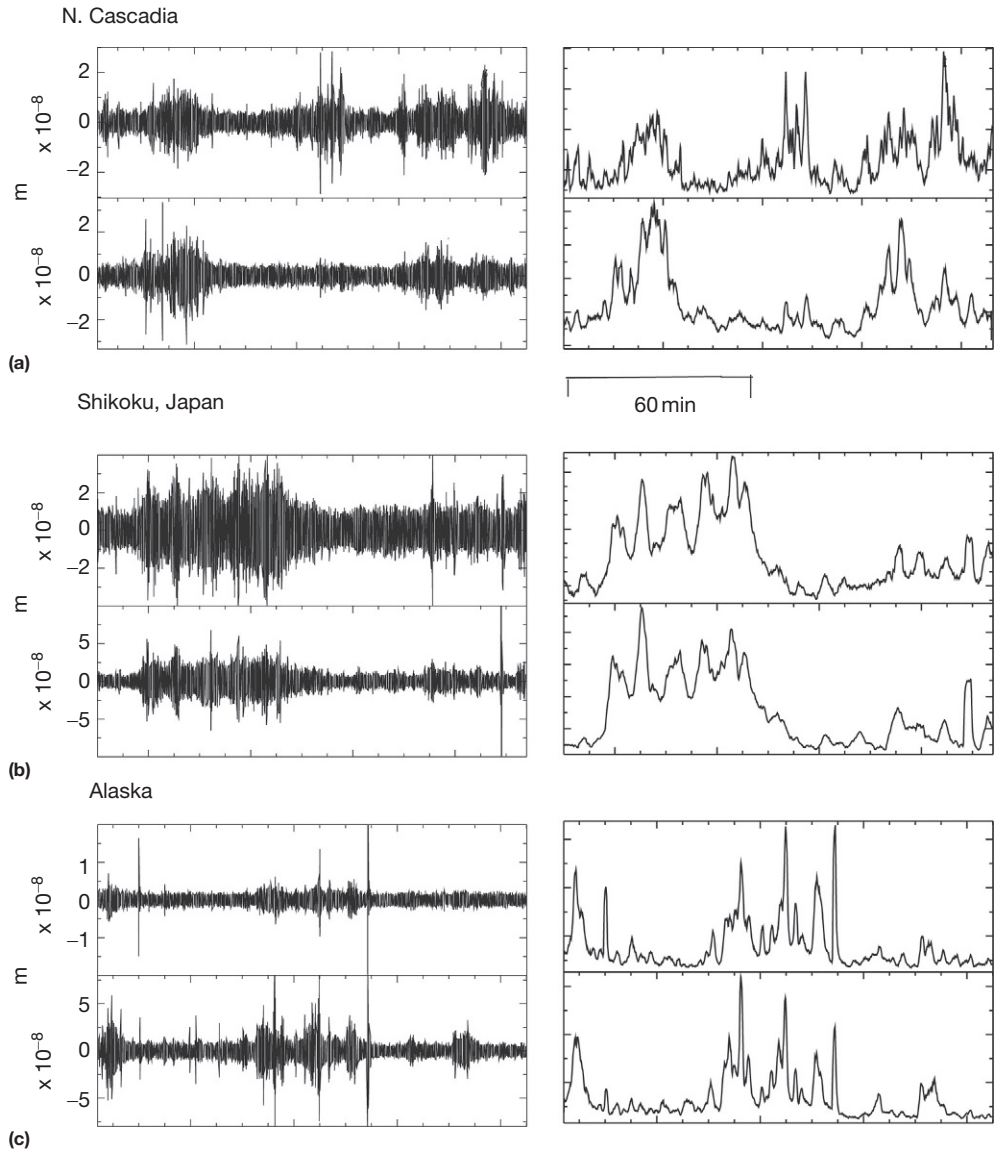


Figure 7 Horizontal records of seismic tremor (left) from two stations in (a) northern Cascadia; (b) western Shikoku, Japan; and (c) southeastern Alaska. All records are 8400 s long and bandpass-filtered between 2 and 4 Hz. Note that amplitude scales vary between records. Right: Envelope-processed versions of the same records. These envelopes emphasize periods of tremor that can be correlated between stations for identification and location.

However, the recording of Cascadia seismic tremor by dense seismic arrays (Ghosh et al., 2010) and the use of improved tremor detection methods (Bostock et al., 2012; Brown et al., 2009) has revealed that it too consists of swarms of LFEs with complex migration patterns that propagate with slow slip at velocities on the order of kilometers per day. The term episodic tremor and slip (ETS) was coined to describe the close relationship between SSEs and tremor in Cascadia and is also used in SW Japan. The close association between slow slip and tremor in these two regions has allowed estimates of the size of SSEs to be made based on the duration of tremor during the SSE (Aguiar et al., 2009; Hirose et al., 2010).

Tremor with a character similar to that recorded in subduction zones has been observed on the Parkfield and Cholame (Nadeau and Dolenc, 2005; Shelly, 2010) segments of the San Andreas Fault. Although the mechanisms behind subduction and continental fault tremor are not necessarily identical, the similarity between the two phenomena argues for some relationship. If tremor is produced by the rupture of small asperities embedded within regions of accelerated aseismic slip (Shelly et al., 2006), then slow slip and tremor may be manifestations of the same phenomena and should be associated with one another regardless of tectonic environment. **Figure 5** identifies regions where seismic tremor is well correlated with SSEs, where both phenomena occur but their association is less clear, and where slow slip occurs but tremor has yet to be reported. For recent reviews of tectonic tremor see Rubinstein et al. (2010), Beroza and Ide (2011), and Ide (2012).

4.17.4 Global Observations of EAS

4.17.4.1 Convergent Margins

SSEs occur most often on the plate interface at convergent margins and exhibit a large variation in duration, magnitude, and recurrence behavior. In fact, almost all subduction zones instrumented with a GPS network sufficiently dense to detect SSEs have done so. One notable exception is the subduction zone offshore of northern Honshu, where a very dense GPS network has recorded several episodes of afterslip, but no interseismic SSEs. Like afterslip, most observed interseismic slow slip locates at the down-dip extension of the seismogenic zone, where the frictional regime may be conditionally stable (**Figure 1(b)**). The large range in behaviors exhibited by SSEs may provide clues to their generation. Documenting the behavior of these events in all subduction zones and comparing and contrasting their behaviors may contribute to a better understanding of how and why they occur and how much they perturb the local stress field to bring the megathrust closer to failure in a great earthquake.

4.17.4.1.1 Cascadia

The Cascadia subduction zone, extending 1000 km from British Columbia to northern California, has a history of generating $M_w > 8.0$ earthquakes approximately every 300–500 years, with the last great earthquake occurring in 1700 (Satake et al., 2003). Geodetic measurements made over the last 15 years have shown that the plate interface along this entire segment is locked from near the surface to a depth of about 20 km (Dragert et al., 1994). This region of geodetic locking has been identified as the seismogenic zone. One of the earliest observations of a subduction zone SSE was made along the northern portion of the Cascadia margin from continuous GPS data (Dragert et al., 2001). These data revealed that in the fall of 1999, the contractional motions indicative of strain accumulation at seven sites in southern Vancouver Island and western Washington suddenly reversed their motion for a period of about 2 weeks at each station (**Figure 4(b)** and **4(c)**). Dragert et al. (2001) found that the surface displacement data associated with this excursion could be well fit by ~2 cm of slow slip on a deep portion of the plate interface just below the seismogenic zone. This SSE had moment equivalent to an $M_w = 6.7$ earthquake. Following this discovery, careful analysis of the continuous GPS time series from this network dating back to 1992 revealed the occurrence of eight similar SSEs with a recurrence interval of 13–16 months (Miller et al., 2002). Each SSE lasted between 2 and 4 weeks at a particular station, migrated across the array, and was consistent with the slip model obtained for the 1999 event, that is, a few centimeters of slip on the deep (25–45 km) portion of the plate interface, yielding equivalent moment magnitudes of 6.5–6.8 (Dragert et al., 2004). Inversion of surface displacements from the 1999, 2003, and 2008 Cascadia SSEs for models of the spatial and temporal variations in fault slip (Bartlow et al., 2011; McGuire and Segall, 2003; Melbourne et al., 2005) indicate that, similar to large earthquakes, SSEs possess rupture complexity and concentrate their slip at isolated patches that form asperities along the fault plane.

Rogers and Dragert (2003) demonstrated a strong correlation between six Cascadia SSEs and seismic tremor (**Figure 6(a)**). Because of the correspondence and regular recurrence of tremor and SSEs in northern Cascadia, Rogers and Dragert (2003) termed this coupled phenomena episodic tremor and slip or ETS. Tremor and slip appear to migrate together in time, although the locations and migration patterns vary between episodes. Migration speeds vary between 5 and 15 km day^{-1} along strike during slow slip episodes. High-resolution studies of seismic tremor reveal plate motion parallel migration of tremor sources with velocities up to tens of kilometers per hour can accompany along-strike slower propagation (Ghosh et al., 2009).

Continuous GPS and seismic data from northern California, Oregon, and southern Washington suggest that ETS occurs along the entire Cascadia margin, but has different characteristics in the different locations. For northern California, Szeliga et al. (2004) detected eight SSEs between 1997 and 2003 based on GPS position changes from a single station; the most recent five SSEs were accompanied by seismic tremor recorded by seismic stations of the Northern California Seismic Network. The average value of the horizontal surface displacement during these events (4–6 mm) is similar to the northern Cascadia SSEs; however, their recurrence interval is 9–12 months, shorter than the 13–16 months obtained in northern Cascadia. Although GPS station coverage is less dense along the southern compared to northern Cascadia margin, several GPS stations along this section show sudden reversals of motion that have been identified as SSEs (Boyarko and Brudzinski, 2010). In most cases, the sudden GPS motion reversals can be correlated with seismic tremor episodes, enhancing the evidence that they represent slow slip on the subduction plate interface (Boyarko and Brudzinski, 2010; Brudzinski and Allen, 2007; Szeliga et al., 2004). In summary, the entire Cascadia margin appears to generate ETS with recurrence intervals that vary with location along the margin. There appears to be three ~300–500 km, coherent zones of ETS progressing from northern Vancouver Island to northern California with recurrence intervals of approximately 14, 19, and 10 months, respectively. The spatial extent of these regions may correlate with upper plate geologic terranes (Brudzinski and Allen, 2007).

4.17.4.1.2 Southwest Japan

Southwestern Japan is a complex region of subduction where the Philippine Sea plate subducts beneath the Eurasian or Amurian plate along the Nankai and Sugura Troughs and beneath the North American plate along the Sagami Trough. Great earthquakes and SSEs occur along most of this convergent margin. SSEs have been detected from network GPS observations in the Bungo Channel, Tokai, and Boso regions. The operation of a very dense network of borehole tiltmeters in Japan since about 2000 (National Research Institute of Earth Science and Disaster Prevention or NIED Hi-net array) has allowed very small SSEs, not detected on surface GPS stations, to also be identified. Although these tilt changes are too small to be recognized as significant on their own, their strong correlation and migration with seismic tremor (**Figure 6(b)**) indicate that they are related phenomena and by association with ETS at Cascadia interpreted to reflect slow slip on the plate interface (Obara et al., 2004). The SSEs detected by GPS typically last months to years, have equivalent moment magnitudes greater than 6.5 and recurrence intervals of more than 5

years, and are referred to as long-term SSEs. Short-term SSEs, detected with tiltmeters, are smaller, last only several days to a week, and reoccur every 3–12 months (Hirose and Obara, 2005, 2006; Obara and Hirose, 2006; Obara et al., 2004; Sekine et al., 2010).

The first GPS deformation signal to be interpreted as an SSE was recorded by the Japanese nationwide GPS network GEONET between 1996 and 1998 in the Bungo Channel region of southwestern Japan (Hirose et al., 1999). A sudden reversal of surface displacements at several stations in the vicinity of the Bungo Channel, initiating in early 1997 and lasting ~300 days, suggested slow slip on the plate interface. Hirose et al. (1999) modeled the surface displacement pattern and found that it was best fit by variable slip down-dip of the seismogenic zone (Hyndman et al., 1995). This model yielded an equivalent earthquake moment magnitude of 6.6. Similar slow slip motions recurred in 2003 (Ozawa et al., 2004) and 2010 (Hirose et al., 2010). Ozawa et al. (2004) inverted the 2003 GPS observations for the spatial and temporal distribution of slow slip on the fault plane and found a slip distribution very similar to that of the 1997 SSE and an equivalent earthquake magnitude of $M_w = 7.0$. During both events, slip concentrated at the transition in plate coupling from strongly coupled in the northeast to weakly coupled in the southwest (Ito and Hashimoto, 2004; Ito et al., 1999).

Continuous GPS data from GEONET stations in the Tokai region revealed a reversal in direction and change in deformation rate beginning in mid 2000 that persisted into 2006. These motions were shown to be consistent with slow slip on a deep (25–40 km) portion of the plate interface (Miyazaki et al., 2006; Ohta et al., 2004; Ozawa et al., 2002). Miyazaki et al. (2006) inverted the GPS data in the time range between January 2000 and November 2002 for the spatiotemporal distribution of slow slip on the fault plane. Their results indicated two slow subevents with slip in each that accelerated and decelerated several times but concentrated down-dip of the portion of the fault plane previously determined to be strongly coupled (Ohta et al., 2004). The maximum slip attained in this time period was ~14 cm equivalent to an earthquake of $M_w = 6.8$; however, the SSE continued for several more years eventually attaining an $M_w = 7.1$ (Suito and Ozawa, 2009). Analysis of tilt observations at a single station in the region between 1981 and the present revealed anomalous tilt associated with the recent Tokai SSE as well as a nearly identical anomaly between 1988 and 1990 (Yamamoto et al., 2005). Yamamoto et al. (2005) interpreted the 1988–90 signal as an SSE and speculated that SSEs in the Tokai region occur repeatedly; however, they noted that the size and duration of the recent SSE far exceed those of its predecessor and that variations in SSEs in this region must exist.

Continuous GPS data between May and June 1996 revealed anomalous horizontal motions on the eastern Boso Peninsula that were interpreted as an SSE (Sagiya, 2004). Inversion of this data for the distribution of the slow slip indicated that approximately 50 mm of slip occurred at shallow depth (10–20 km) on the fault plane within a period of ~1 week, yielding an equivalent moment magnitude of 6.4. The source region of the SSE was in the same depth range as interseismic locking (Sagiya, 2004) and the 1923 Kanto earthquake asperity (Wald and Somerville, 1995), albeit in a different location along strike, farther to the east. This is unusual in that all the other SSEs in southwest Japan have been located down-dip of regions of strong interseismic locking and coseismic rupture. The source area of this SSE is at the boundary between a strongly locked and more freely slipping region of the plate interface, which corresponds to a large contortion in the downgoing Philippine Sea plate, perhaps because of its interaction with the westward subducting Pacific Plate. A second SSE was documented in this region over a 50-day period beginning in October 2002 (Ozawa et al., 2003). Ozawa et al. (2003) inverted the GPS data to obtain the time evolution of the 2002 Boso event and to compare it with characteristics of the SSE 6 years previously. They found that the magnitude of the 2002 event ($M_w = 6.6$) exceeded the 1996 event and although their slip areas overlapped, the centroid of slip during the 2002 event was farther south. A third SSE in this area was recorded in 2007 with very similar parameters to the previous events (Ozawa et al., 2007). The Boso SSEs are different from the Bongo Channel and Tokai geodetically detected events in that they are much shorter in duration and occur at shallow depth.

Many episodes of short-term slow slip have been observed in western Shikoku, Kii, and Tokai between 2001 and now (see Obara, 2010, 2011 for reviews). Like Cascadia, these SSEs are accompanied by seismic tremor and the spatiotemporal characteristics of tremor and slip divide this portion of the margin into segments with regular recurrence intervals (Sekine et al., 2010). Dislocation modeling of the tilt data at several stations has allowed source parameters for some of these to be determined (Hirose and Obara, 2005; Obara and Hirose, 2006; Obara et al., 2004), revealing that slip occurs at the down-dip edge of the plate interface at depths between 20 and 45 km, a bit deeper than the long-term SSEs. Their source areas appear to define several patches on the deep plate interface, with each SSE activating a slightly different combination of these patches (Hirose and Obara, 2005). The average slip per event is between 1 and 4 cm, yielding equivalent moment magnitudes between 5.5 and 6.2. While the long-term SSEs seem to repeat every 5–6 years, the recurrence interval for short-term SSEs is 2–6 months.

Biannually repeating SSEs have been detected along the southern Ryukyu Arc southwest of the Nankai Trough (Heki and Kataoka, 2008). These events appear to have displacements of about 5–6 cm (M_w 6.6) along the same patch of the plate interface at a depth of 20–40 km. Unlike the SSEs that occur regularly at the Nankai Trough, the Ryukyu events occur along an unlocked segment of the plate interface that has never generated a large underthrusting earthquake. Heki and Kataoka (2008) suggested repeated rupture of a former asperity that transitioned from fast to slow slip because of an increase in temperature.

4.17.4.1.3 Mexico

Subduction of the Cocos plate beneath the North American plate along the Middle American Trench in Mexico generates large to great earthquakes every 30–100 years. The Guerrero and Oaxaca regions of southern Mexico also experience regular SSEs. The last great earthquake that occurred along the Guerrero segment in southern Mexico was in 1911, which makes it a likely location for rupture in a future large earthquake. Data from a single continuous GPS station in the Guerrero region

recorded transient surface displacements in early 1998 that were interpreted as slow slip on the plate interface (Lowry et al., 2001). The continuous GPS data from this station, campaign data from several other stations in the area, and leveling data were consistent with moment approximately equivalent to an $M_w > 6.5$ earthquake that propagated along strike over several months. In 2001, seven permanent GPS stations in the Guerrero region recorded a reversal of surface motion consistent with an SSE beginning in October 2001, lasting approximately 7 months with an $M_w = 7.5$ (Kostoglodov et al., 2003). This is the largest SSE that has been recorded globally, and continued GPS monitoring has revealed that similar events repeated in 2006 (Larson et al., 2007), and 2009/10 (Walpersdorf et al., 2011), yielding a recurrence interval of approximately 4 years. Campaign and continuous GPS data collected in the Guerrero region between 1992 and early 2001 were inverted to obtain both the interseismic deformation pattern and the location of transient slip events during this period. The pattern of interseismic strain accumulation indicated a locked plate interface at depths shallower than 25 km. Modeling of reversals of this motion indicated that SSEs in 1995–96 ($M_w = 6.8$–7.1) and in 1998 ($M_w = 6.8$–7.1) occurred at depths greater than 25 km (Larson et al., 2004). Modeling of GPS data for the last three large SSEs in the Guerrero region (Radiguet et al., 2011) found similar slip distributions, with maximum slip of ~15 cm and slip extending to a shallower depth, possibly into the seismogenic zone (15–20 km, depending on the slab model used). The regions of slow slip correlate with the existence of an ultra-slow velocity layer at the top of the subducting oceanic crust that is inferred to represent a zone of high pore fluid pressure (Song et al., 2009). SSEs were also reported in the Oaxaca region of Mexico in 2004 (Brudzinski et al., 2007), 2006, and 2007 (Correa-Mora et al., 2008, 2009). All these events occurred at depths below the seismogenic zone, lasted about 3 months, and had equivalent moment magnitudes of 7.0–7.3 (Correa-Mora et al., 2009). Although the 2006–07 slow slip activities in Guerrero and Oaxaca occurred close in time, their source regions were separated by over 100 km. The existence of two independent slow slip areas in southern Mexico is supported by the different patterns of surface deformation; larger slow slip displacements occur along the coast in Guerrero compared with Oaxaca. This is consistent with Guerrero experiencing a shallower and larger magnitude transient slip than Oaxaca (Vergnolle et al., 2010). Seismic tremor has been observed in both Guerrero and Oaxaca; however, it is more frequent, shorter in duration, and located further inland than GPS-detected slow slip (Brudzinski et al., 2010; Payero et al., 2008).

4.17.4.1.4 New Zealand

Since 2002, more than 15 SSEs of variable size and duration have been documented along the Hikurangi margin based on continuous GPS records (Wallace and Beavan, 2010; Wallace et al., 2012). A reversal of surface displacements over a 10-day period in October 2002 at two continuous GPS stations on the northeast coast of New Zealand's North Island was the first reported SSE in this region (Douglas et al., 2005). The surface displacements were modeled by 18 cm of slip on the plate interface at the down-dip edge of the seismogenic zone. With the expansion of the continuous GPS network, a bimodal distribution of SSEs has been recognized. Large ($M_w \sim 7.0$), relatively deep (20–60 km), long duration (1–1.5 years), and recurrence interval (~5 years) SSEs occur in the central and southern Hikurangi subduction zone while smaller ($M_w \sim 6.5$–6.8), shallower (<15 km), shorter duration (1–3 weeks), and recurrence interval events (1–2 years) occur along the east coast in the central and northern part of the margin (Wallace and Beavan, 2010). Long-term versus short-term SSE behavior may be controlled by the heterogeneity and/or effective normal stress on the plate interface, with increased heterogeneity and lower normal stress favoring short-term SSEs (Wallace et al., 2012). Heterogeneity in frictional properties of the plate boundary may also be responsible for the patchy migration of slow slip at this margin compared with Cascadia (Wallace et al., 2012). Tremor has been observed to accompany short-term SSEs in the Gisborne region (Kim et al., 2011) and long-term SSEs farther south (Ide, 2012), but occurs for much less time and in shorter duration bursts than at most other subduction zones.

4.17.4.1.5 Costa Rica

Three continuously recording GPS stations on the Nicoya Peninsula, Costa Rica, recorded a transient deformation event lasting approximately 1 month in September–October 2003. The transient displacements were nearly opposite in direction to plate convergence and strain accumulation on the plate interface, and appeared to have initiated at the up-dip transition from a locked to a more freely slipping interface and propagated landward (downdip). The transient displacements were interpreted as an underthrusting slow earthquake located primarily within the seismogenic zone (Jiang et al., 2012; Protti et al., 2004). Three SSEs in early 2000, lasting approximately 3 weeks each, were proposed in the Nicoya region, based on observations of correlated transient fluid flow and seismic tremor (Brown et al., 2005). LaBonte et al. (2009) used a 2-D fully-coupled poroelastic finite-element model to show that the patterns of fluid flow could be matched by a dislocation on the plate interface at a depth of less than 4 km, propagating bilaterally and lasting 20 days or more. Since these initial observations, four additional significant SSEs have been detected with GPS data, yielding a 2-year recurrence interval (Jiang et al., 2012). Minor SSEs visible on a few GPS stations and accompanied by active tremor (Walter et al., 2011) occur more frequently, suggesting that this mode of strain release is common along the Nicoya segment of the Costa Rica subduction zone. Slip in six of the eight well-recorded SSEs ($6.1 < M_w < 6.9$) locates off shore the Nicoya Peninsula at shallow depth (<15 km) on the plate interface (Jiang, 2012). Shallow slow slip has now been documented at three subduction zones: Boso, Japan; northern Hikurangi, New Zealand; and Nicoya Peninsula, Costa Rica.

4.17.4.1.6 Alaska

The eastern end of the Alaska–Aleutian subduction zone is the site of the great M_w 9.2 1964 Prince William Sound earthquake. This event ruptured a large region of the shallow plate interface above 30 km depth and had a long and complicated postseismic deformation history (Cohen and Freymueller, 2004). Ohta et al. (2006) inverted GPS data for the period between 1997 and 2002 and found two distinct deformation patterns: while data before 1998 and after 2001 showed surface

displacements consistent with strain accumulation over the entire shallow plate interface, in contrast, data between 1998 and 2001 indicated continued strain accumulation on the shallow plate interface and an SSE on the deeper interface in an isolated region north of Anchorage. The region with the largest slip deficit or maximum locking lies beneath the southeast Kenai Peninsula, co-located with the asperity that broke in the 1964 earthquake. During the period between 1998 and 2001, the surface velocities north of Anchorage differ from those in the other time periods, which is consistent with the occurrence of an SSE. The best-fitting fault parameters for this SSE were found to be maximum and average slip rates of 55 and 40 mm year^{-1} respectively, producing total slip of 12.0–16.5 cm over a fault area of $\sim 150 \times 150$ km^2 for the 3-year duration of the event. The cumulative seismic moment was 1.1×10^{20} Nm, which corresponds to M_w 7.2 (Ohta et al., 2006). The maximum slip area of the SSE locates at depths between 25 and 45 km, just below the seismogenic zone. Peterson and Christensen (2009) reported seismic tremor at stations spanning over 100 km along the Alaska/Aleutians trench over several months each year of the SSE. Another SSE was detected in early 2010 and continued through at least the end of 2011 (Wei et al., 2012). This event had an M_w of $\sim$6.9 and also occurred at the down-dip edge of the 1964 earthquake rupture area, but farther to the southwest compared to the 1998–2001 SSE (Wei et al., 2012).

4.17.4.1.7 Northeast Japan

Strain release along the Pacific–North American plate interface in northeast Japan occurs as great earthquakes, many of which have experienced significant amounts of afterslip, and continuous aseismic slip with no documented observations of interseismic SSEs. This is noteworthy because of the presence of dense networks of continuous GPS and tiltmeter stations in this region that are comparable to those in southwest Japan, where SSEs occur frequently. Suwa et al. (2006) inverted both the horizontal and vertical components of the GPS GEONET data to obtain the pattern of interseismic strain accumulation on the plate interface in northeastern Japan. They found a few patches of strong locking that extended to about 60 km in depth beneath north-central Honshu and to a greater depth beneath Hokkaido. A comparison of this pattern of interplate coupling with the asperity distribution for all earthquakes with $M_w > 7.5$ in this region determined by Yamanaka and Kikuchi (2003, 2004) reveals a very close correspondence. This indicates that the regions of maximum coseismic slip during large earthquakes heal quickly to become the loci of strain accumulation in the interseismic period. Significant afterslip that followed the 1994 Sanriku-Oki ($M_w = 7.7$) and 2003 Tokachi-Oki ($M_w = 8.0$) earthquakes was found to primarily occur in regions that slip during interseismic periods. Afterslip following smaller earthquakes in this region has also been detected by strainmeters for the 1989 ($M_w = 7.4$) and 1992 ($M_w = 6.9$) Sanriku-Oki earthquakes (Kawasaki et al., 1995, 2001), by tide gauge and leveling for the 1978 Miyagi-Oki ($M_w = 7.4$) earthquake (Ueda et al., 2001), and by GPS for the moderate-sized 2001 ($M_w = 6.4$) Tokachi-Oki earthquake (Sato et al., 2004). It appears that afterslip in the weakly coupled regions of the plate interface is a common occurrence following moderate to great earthquakes. The weakly coupled portions of the fault plane that host earthquake afterslip also generate a large number of small repeating earthquakes (Igarashi et al., 2003). Repeating earthquakes have nearly identical waveforms recorded at the same stations, indicating that they rupture similar fault locations. Repeating earthquakes were first observed along the San Andreas Fault system in California and were interpreted as the repeated rupture of small isolated asperities embedded within freely slipping regions (Nadeau and Johnson, 1998). Uchida et al. (2003) contoured the distribution of repeating earthquakes on the plate boundary in northeastern Japan and found several concentrations of repeating earthquakes that they interpreted as freely slipping portions of the plate boundary. A comparison of the repeating earthquake clusters with the geodetically determined distribution of plate coupling confirms that they do occur where the plate interface is freely slipping. More recently, two sequences of migrating foreshocks of the great 2011 Tohoku-Oki (M_w 9) earthquake have been interpreted as slow-slip transients propagating toward and leading up to rupture initiation (Kato et al., 2012). Stress loading from the second sequence, which involved large slip rates, has been identified as the probable trigger of unstable dynamic rupture of the mainshock. In summary, the plate interface in northeastern Japan appears to be composed of (1) regions that are accumulating strain at rates close to plate convergence that will become the loci of future large earthquakes; (2) regions that are continuously slipping at or near plate rates; and (3) regions that slip at intermediate rates somewhere between those of earthquakes and plate rates, but only when driven by coseismic slip in adjacent regions (these are regions of afterslip). Although several different modes of strain release are observed, there is no evidence that EAS-like events occur during the interseismic period in northeast Japan.

4.17.4.2 San Andreas Fault

Aseismic slip is a common mode of strain release within the San Andreas Fault system and is well documented along the Maacama, Concord, Hayward, Calaveras, and central San Andreas faults (Galehouse, 2002). This slip occurs as shallow subsurface creep and at seismogenic depth and includes both steady-state and episodic motions. The episodic shallow creep appears to be caused primarily by soil failure (Johnston and Linde, 2002 and references therein) and is not discussed further in this chapter. Aseismic strain transients at seismogenic depth have been reported as slow-slip episodes, inferred from geodetically detected rate variations (Gao et al., 2000; Langbein et al., 1999), characteristics of repeating microearthquakes (Nadeau and McEvilly, 1999, 2004), and SSEs inferred from borehole strainmeter data (Linde et al., 1996). The locations of these EAS events are indicated in **Figure 8**. Several SSEs were reported along the central San Andreas Fault near San Juan Bautista in 1992 (Linde et al., 1996), 1996 (Johnston, 1997; Johnston et al., 1996), 1998 (Gwyther et al., 2000), 2003, and 2004 (Gladwin, 2004; Pollitz and Johnston, 2006); however, only the first of these is well documented. The December 1992 event was recorded by three borehole strainmeters and two surface creepmeters located near the fault. The data (e.g., at station SJT in **Figure 4(d)**) revealed a complicated series of strain changes that began abruptly, contained many subevents, and lasted for approximately 5 days.

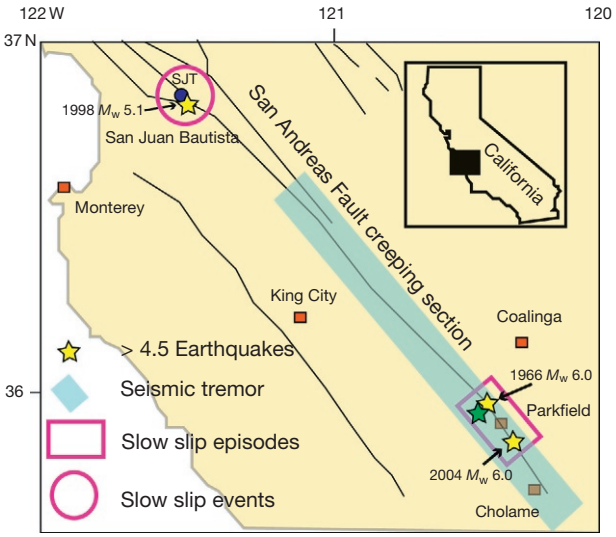


Figure 8 Map of central coastal California indicating the location of EAS associated with the creeping section of the San Andreas Fault. The large circle and rectangle locate regions of SSEs near San Juan Bautista and changes in aseismic slip rates near Parkfield. The blue shaded region indicates the extent of seismic tremor, the yellow stars mark the location of significant large earthquakes, and the green star shows the location of $\sim M$ 4.5 earthquakes that preceded slip rate changes at Parkfield. The circle labeled SJT shows the location of the borehole strainmeter for which data are shown in **Figure 4**.

Forward modeling of these data indicated a slip sequence with complexities similar to those observed during normal earthquakes. The source processes consisted of several slip episodes scattered in the upper 8 km. The $\sim$50 km^2 fault plane had slip amplitudes and durations ranging from 0.3 to 2.5 cm and 0.6 to 43 h, respectively (Linde et al., 1996). The total moment of this sequence was equivalent to an M_w 5 earthquake.

Episodic slow slip at Parkfield, CA, was identified by abrupt changes in electronic distance meter (EDM) line length measurements, tensor strain, and a 30% increase in slip rate on the fault inferred from the recurrence interval of repeating earthquakes beginning in early 1993 and lasting until 1996 (Gao et al., 2000; Langbein et al., 1999; Nadeau and McEvilly, 1999). The deviations in surface deformation were successfully modeled as acceleration in slip rate at seismogenic depth on the San Andreas Fault (Gao et al., 2000; Gwyther et al., 1996; Langbein et al., 1999). While previous studies assumed constant slip rates over fixed time intervals, Murray and Segall (2005) analyzed EDM data over a longer time period (mid-1980s to 2003) to image the time-dependent evolution of slip on the fault. They found episodic slip rate changes between 1993 and 1996, with the largest increases correlating with the occurrence of three M $\sim$4.5 earthquakes (**Figure 8**), the largest events to occur in the Parkfield region between 1966 and 2004. Slip was constrained to occur in the upper 8 km of the crust with a maximum estimated slip rate of 4.9 cm year^{-1}, compared with the long-term slip rate of $\sim$3.5 cm^{-1} year^{-1}. Nadeau and McEvilly (2004) determined slip rates as a function of time along the entire 175-km creeping section of the central San Andreas Fault (**Figure 8**) from analysis of repeating microearthquakes. They found considerable variability, with rate changes as large as 100% throughout the region, and some areas that possessed quasi-periodic rate changes with durations and recurrence periods of 1.7–3 and 2.5–3 years, respectively. Nadeau and McEvilly (2004) favorably compared these slip episodes to interseismic SSEs at convergent margins, concluding that both phenomena may share a common mechanism. No seismic tremor has been associated with the San Juan Bautista SSE or episodic slip on the central San Andreas Fault (SAF). However, deep seismic tremor (20–30 km depth) occurs frequently along the San Andreas Fault from the Cholame section, some 40 km south of Parkfield (Nadeau and Dolenc, 2005; Shelly, 2010) northwest to Bittersweet, $\sim$ 80 km north of Parkfield (Shelly and Hardebeck, 2010; **Figure 8**). If slow slip can be inferred from observations of deep tremor, as has been done in several subduction zone settings, the widespread tremor distribution implies that slow slip occurs along the entire creeping section of the San Andreas Fault. Shelly (2009) detected possibly anomalous tremor migration accompanied by elevated tremor rates in the 3 months prior to the 2004 M_w 6.0 Parkfield earthquake and suggested precursory slow slip at comparable depth to the eventual earthquake hypocenter.

4.17.4.3 Hawaii

A continuous GPS network on Kilauea volcano, Hawaii, recorded transient displacements in November 2000 that were interpreted as an SSE (Cervelli et al., 2002). Inversion of the GPS data indicated that slip occurred on a shallow dipping thrust fault at a depth of $\sim$5 km beneath the southeast flank of the volcano. The slip attained a maximum rate of 6 cm day^{-1}, and lasted for about 36 h, producing 8.7 cm of total slip with an equivalent moment magnitude of 5.7 (Cervelli et al., 2002). A peak in rainfall 9 days before this SSE suggests a causal relationship. Cervelli et al. (2002) postulated that an elevated water table, resulting from nearly 1 m of rainfall in a short period of time, increased the pore pressure and decreased the effective normal stress on crustal faults, triggering slip. Examination of GPS data from Kilauea through 2007 revealed seven additional SSEs (Brooks et al., 2006; Montgomery-Brown et al., 2009), most not associated with periods of increased rainfall. The four largest SSEs ($M_w \sim 6.0$) appear periodic with a recurrence interval of about 2 years (Brooks et al., 2006; Montgomery-Brown et al., 2009). Four smaller events have also been identified and, if included, yield an SSE recurrence interval of just less than 1 year (Montgomery-Brown et al., 2009). Source parameters for all SSEs are consistent with motion on a shallow reverse fault, similar to the model proposed by Cervelli et al. (2002), with comparable seismic moments. Although most Kilauea SSEs are accompanied by an increase in microseismicity (Brooks et al., 2006; Segall et al., 2006), so far no tectonic tremor has been identified at Kilauea.

4.17.4.4 Oceanic Transform Faults

EAS has been invoked to explain anomalous seismic or geodetic behavior observed in several other tectonic settings as well. The rupture process of large oceanic transform faults has been shown to often involve slow precursory slip followed by compound rupture on spatially distinct faults (e.g., McGuire et al., 1996; McGuire and Jordan, 2000). Support for this idea

comes from observations of transform fault microseismicity. Forsyth et al. (2003) reported alternating slip between adjacent transform faults during an earthquake swarm recorded by Ocean Bottom Seismometers in the southern Pacific Ocean. Because static or dynamic triggering by such small earthquakes over relatively large distances (25 km) was deemed unlikely, Forsyth et al. (2003) suggested EAS to explain event coupling during these earthquake swarms. Observations of high foreshock to aftershock rates for events along the East Pacific Rise, compared to continental strike-slip faults, prompted McGuire et al. (2005) to invoke slow slip transients as a mechanism for triggering these earthquakes. Similarly, Vidale and Shearer (2006) found that over 70 earthquake swarms in southern California did not follow typical foreshock–mainshock–aftershock behavior and concluded that EAS was the most probable explanation for the observed steady rate of seismicity.

4.17.5 EAS and Seismicity

Afterslip associated with moderate to large earthquakes suggests a relationship between EAS and higher-velocity rupture. In addition, models of EAS generation that appeal to small perturbations in the stress field (e.g., Shen et al., 2005) imply that minute stress changes related to nearby earthquakes can trigger slow slip. Conversely, one of the motivations behind the increased study of interseismic EAS is the potential triggering effect these events may have for normal earthquakes. Many individual observations suggesting a triggering relationship between EAS and seismic behavior have been made including (1) the occurrence of moderate to large earthquakes associated with EAS; (2) changes in seismicity rates preceding or following EAS; and (3) changes in slip rates following moderate to large earthquakes; however, calibrating the pattern requires more work.

The most extensive work on seismicity patterns and EAS was performed for the multi-year Tokai SSE of 2000. Several authors reported a relative quiescence in shallow (<50 km) seismicity prior to slip in Tokai (Matsumura, 2006; Yamamoto et al., 2005). During the SSE, Yoshida et al. (2006) reported variations in slab and crustal seismicity concurrent with changes in slip velocity. The rate of slab earthquakes ($M > 1.1$) increased at the initiation of slow slip, decreased during a period of slower slip, and rose again, coincident with an increase in slip. The fact that this correlation is made using seismicity down to $M = 1.1$ suggests that a highly complete catalog may be necessary for tracking any subtle changes in seismicity resulting from slow slip.

Hirose et al. (1999) suggested that the 1996–97 Bungo Channel SSE was affected by changes in stress caused by the occurrence of two large M_w 6.7 earthquakes near Hyuganada, 100 km to the southwest, about 3 months prior to the SSE. These two large earthquakes were accompanied by significant afterslip. Although Miyazaki et al. (2003) showed that the regions of afterslip and slow slip did not overlap, they may still have been related through static stress transfer. It is possible that stress perturbation caused by the Hyuganada events contributed to the longer duration and larger moment of the 1996–97 compared with the 2003 SSE, but this possibility is difficult to constrain. In addition, a small swarm of earthquakes (maximum $M_j = 4.9$) occurred north of the Bungo Channel in April 1997 just after the initiation of slow slip (Hirose et al., 1999; Miyazaki et al., 2003).

Ozawa et al. (2003) noted a seismic swarm off the coast of Boso Peninsula on the upper surface of the Philippine Sea plate (20–30 km) concurrent with the region's 2002 SSE. A similar swarm occurred during the 1996 and 2007 SSEs (Ozawa et al., 2007), which suggested a consistent relationship between microseismicity and slow slip (Sagiya, 2004). Ozawa et al. (2003) suggested that this swarm activity increased stress in the neighboring region and caused the observed slow slip. Consequently, these authors used other peaks in seismicity to propose additional slip events in 1983 and 1991–92 (consistent with the inferred 6–7 years recurrence interval).

A cursory analysis of seismicity patterns of moderate to large earthquakes ($M > 3$ in southwest Japan and $M > 4$ elsewhere) associated with SSEs at six convergent margins (Cascadia; western Shikoku and Boso, Japan; Guerrero, Mexico; New Zealand, and Costa Rica) revealed a few isolated correlations, but no consistent cause and effect relationship (Schwartz and Rokosky, 2007). The most robust pattern occurred in Costa Rica where Brown et al. (2005) reported a rise in cumulative moment 20 days after each of three correlated fluid flow and seismic tremor episodes that occurred in 2000 and were interpreted as SSEs. A series of earthquakes to the northwest of the 2003 geodetically inferred SSE also produced a precipitous rise in cumulative seismic moment and event number following this slip event (Schwartz and Rokosky, 2007). However, there does seem to be a relationship between SSEs and small earthquakes in several subduction zones. Increases in seismicity were reported associated with SSEs in the northern Hikurangi margin of New Zealand (Delahaye et al., 2009; Kim et al., 2011) and in south central Alaska (Wei et al., 2012). While it was assumed that these slow slip triggered events occurred on the plate interface, five small earthquakes within the subducted slab that occurred during a Cascadia SSE were interpreted as also being triggered by the EAS (Vidale et al., 2011).

Well-documented observations of EAS along the San Andreas Fault have also been associated with anomalous seismic behavior. The 1992 San Juan Bautista SSE was accompanied by several small earthquakes (M_L 3.1–3.7), not typical of seismicity patterns before or after this event (Linde et al., 1996). The largest recent earthquake (M_w 5.1) along the San Juan Bautista segment of the San Andreas Fault (**Figure 8**) immediately preceded the 1998 SSE in this region (Uhrhammer et al., 1999). Nadeau and McEvilly (2004) found a correlation between the occurrence of $M > 3.5$ earthquakes and pulses of accelerated slip that they inferred from repeating microearthquakes. They reported a six- to sevenfold increase in seismic activity near the onset of accelerated slip episodes compared with background rates. They could not determine whether accelerated slip preceded or followed earthquake productivity increases; however, they concluded that the quasi-periodic recurrence of accelerated slip episodes and increased seismicity indicated a triggering relationship. Murray and Segall (2005) also reported a correlation between accelerated slow slip and seismicity during the 1993–96 Parkfield, California, slip episode. They found that two of the largest episodic slip rate increases were preceded by the

occurrence of three $M \sim 4.5$ earthquakes (Figure 8), which suggested that the earthquakes may have triggered slow slip.

In Hawaii, an increase in microseismicity was associated with at least four of the SSEs inferred to occur there, with slow slip clearly preceding the seismicity increase for the largest SSE of 2005 (Segall et al., 2006). Assuming that the increased earthquake activity resulted from static stress changes induced by slow slip, Segall et al. (2006) constrained the 2005 SSE to occur on the same, nearly horizontal, plane as the earthquakes. The authors suggested that triggering of microseismicity by slow slip might be a common behavior that goes undetected in regions with low levels of background seismicity (e.g., Cascadia). Seismic tremor triggered by SSEs is common in subduction zones and as described previously has been observed in Cascadia (Dragert et al., 2004), SW Japan (Obara and Hirose, 2006), Mexico (Brudzinski et al., 2010; Payero et al., 2008), Costa Rica (Outerbridge et al., 2010; Walter et al., 2011), New Zealand (Kim et al., 2011), and Alaska (Peterson and Christensen, 2009). Based on the planar location of relocated LFEs embedded within tremor in southwest Japan, Shelly et al. (2006) demonstrated that tremor represents patches of more accelerated shear failure on the plate interface surrounded and driven by broad regions of slow slip. Therefore, EAS likely triggers earthquakes in the adjacent seismogenic zone and seismic tremor in the slow-slipping areas where slow slip induces a positive change in the stress field.

4.17.6 Mechanics of EAS

EAS at convergent margins, occurring either in the interseismic stage or as afterslip, has primarily been detected at the down-dip edge of the seismogenic zone, with the slip continuing into the deeper conditionally stable portion of the plate interface (Figure 1(b)). Notable exceptions include interseismic SSEs in Boso, Japan, and Costa Rica, afterslip following several large earthquakes in northeastern Japan and the 2004 Sumatra and 2005 Nias, Indonesia, earthquakes, and, possibly, precursory slip before the 2011 Tohoku earthquake. EAS along the SAF also occurs within the seismogenic zone. In these cases, however, EAS occurs in complementary locations along strike or up-dip of strongly locked patches (coseismic asperities) or large earthquake hypocenters. No EAS has been located in strongly locked regions of the plate interface. This observation indicates that different frictional properties control the occurrence of fast versus slow slip. Rate- and state-dependent frictional modeling supports this, showing that SSEs require either spatially or temporally variable friction. Shibazaki and Iio (2003) were able to simulate SSEs by introducing temporal variations in their rate- and state-dependent frictional modeling such that steady state friction behaves as velocity weakening at low slip velocity and velocity strengthening at high slip velocity. Their simulated SSEs had slip velocities of tens of meters per year, propagated horizontally at rates of $\sim$10 km year^{-1}, and persisted for tens of years. Yoshida and Kato (2003) and Liu and Rice (2005) successfully simulated SSEs with spatially variable friction using block-spring models and three-dimensional numerical simulations, respectively. In the block-spring models, slip events occurred only in regions where friction was near the stability limit (Yoshida and Kato, 2003). In the numerical simulations, SSEs appeared spontaneously at the down-dip edge of the seismogenic zone, which Liu and Rice (2005) attributed to the transition from unstable to stable frictional properties. Low effective stress and large slip-weakening distances are other model features that have successfully generated slow slip (Liu and Rice, 2007; Rubin, 2008). These simulated strain transients had slip velocities that were meters to tens of meters per year (1–2 orders of magnitude faster than plate velocities). The resulting slip was confined to the down-dip frictional stability transition and migrated along strike at tens of kilometers per year over tens of years. Only the last characteristic is unlike most observations of SSEs in subduction zones in that it is longer than the observations. Transitions in frictional properties from velocity weakening to velocity strengthening seem to be required to generate SSEs. While we expect this transition at the down-dip edge of seismogenic zones, where most SSEs occur, it appears likely that such transitions also exist within seismogenic zones (Figure 9) and are the loci of SSE nucleation. Another possibility is dilatant stabilization in which faults

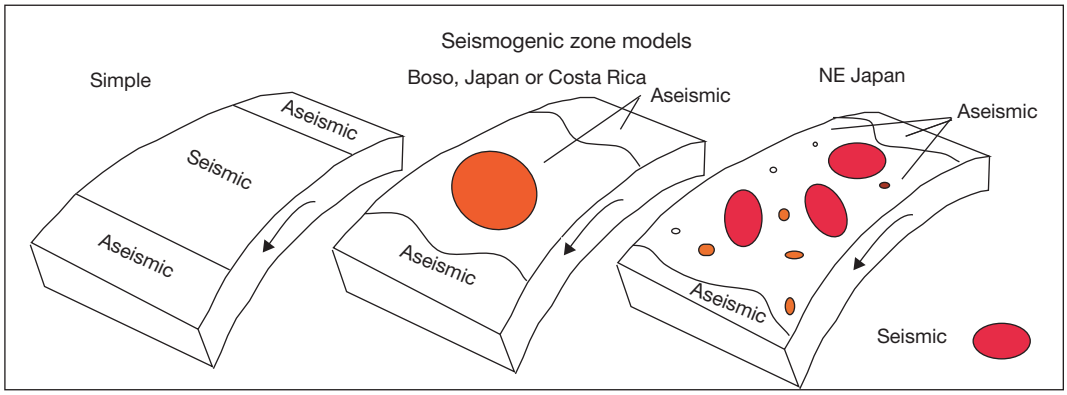


Figure 9 Cartoon of the plate interface at different types of convergent margin seismogenic zones. The simple model contains only two frictional transitions from velocity weakening (seismic)- to conditionally stable or velocity-strengthening (aseismic) behavior at its up- and down-dip limits. The Boso or Costa Rica (middle) and northeast Japan (right) models have these transitions within the seismogenic zone. SSEs nucleate at these frictional transitions, with most occurring at the deeper transition. The greater frictional heterogeneity of the northeast Japan subduction zone may prevent SSEs from initiating within its seismogenic zone.

expand as they slip, allowing pore pressures to drop, raising the effective normal stress, and damping dynamic rupture (Segall et al., 2010). The association of SSEs with areas of high pore-fluid pressure supports this process. As pore-fluid pressure in the fault zone approaches lithostatic, the effect of dilatancy in stabilizing slip is expected to become even more important.

The subduction zones in which little to no interseismic slow slip occurs at or below the seismogenic zone (NE Japan, Boso, and Costa Rica) share some common features that may be important for understanding the generation of slow slip. In particular, the seismogenic zones at these plate boundaries, defined by the 350 °C isotherm (interpreted to represent the transition from velocity-weakening friction to strengthening behavior) and/or the down-dip limit of underthrusting earthquakes or geodetic locking, extend to depths below where the continental Moho intersects the downgoing plate. Three subduction zones that have well-located SSEs at the down-dip edge of the seismogenic zone (Cascadia, SW Japan, and Guerrero, Mexico) have relatively hot plate interfaces in which temperatures reach 350 °C at shallow depths, above where the continental Moho intersects the subducting plates. Although it is unclear which attributes of these hot subduction zones are most important for slow slip generation, reaching the frictional transition at relatively low pressure or above where the oceanic crust of the downgoing plate intersects the Moho of the overlying plate appears to be required for deep slow slip.

The Boso and Costa Rica subduction zones have interseismic SSEs that locate at relatively shallow depths within their seismogenic zones. At Boso, slow slip occurs at depths between 10 and 20 km adjacent to a locked patch, believed to be the likely location of the future Kanto earthquake (Sagiya et al., 1995). In Costa Rica, it occurs at depths between 10 and 25 km, in regions believed to be freely slipping adjacent to strongly coupled patches of the plate interface (Walter et al., 2011). As mentioned earlier, slow slip nucleation most likely occurs at frictional transitions within the seismogenic zone at depths above where the continental Moho intersects the downgoing plate.

In northeast Japan, afterslip is common following great, large-, and even moderate-sized earthquakes, and locates both within and below the seismogenic zone. However, interseismic SSEs have not been identified anywhere along this margin. As discussed previously, afterslip is believed to take place in velocity-strengthening regions, driven by stress concentrations that remain near frictional transitions after earthquakes. Afterslip occurs within the seismogenic zone of northeast Japan indicating the presence of such frictional transitions, so why does not slow slip nucleate at these shallow transitions? Previous work has established that the seismogenic zone in northeast Japan is extremely heterogeneous in its frictional properties, consisting of both large and small patches with velocity-weakening behavior embedded within regions with velocity-strengthening behavior (e.g., Igarashi et al., 2003; Uchida et al., 2003). Perhaps this extreme degree of frictional heterogeneity prevents expanses of velocity-strengthening material large enough to sustain detectable significant slow slip from existing. Numerical simulations by Kuroki et al. (2004) showed that SSEs had large nucleation zones, lending support to this idea. **Figure 9** illustrates this concept, showing seismogenic zones with increasing levels of frictional heterogeneity. The large regions of velocity-strengthening material (labeled aseismic in **Figure 9**) postulated to exist within the Boso and Costa Rica seismogenic zones might host slow slip, while comparably large regions are absent within the northeast Japan seismogenic zone. It is also plausible that the large amount of afterslip following earthquakes in northeast Japan relieves nearly all the stress in the velocity-strengthening regions of the plate interface, leaving little to be released in SSEs.

4.17.7 Recurrence of EAS

Most SSEs are recurrent, with repeat intervals that range from many years (long-term SSEs in the Bungo Channel, Boso, and Tokai, Japan) to a few months (short-term SSEs in Shikoku and Kii Peninsula, Japan, or New Zealand). The recurrence interval for SSEs in northern Cascadia is well established at 13–16 months. Miller et al. (2002) noted that this repeat interval encompasses the 14.5-month periodicity of Earth's Chandler wobble (wobble of Earth's rotation axis caused by mass redistribution at Earth's surface) but dismissed a causative relationship because of the small size of the stress perturbations generated. Shen et al. (2005) quantitatively investigated this relationship by calculating the stress changes induced by the Chandler wobble at the down-dip plate interface in northern and southern Cascadia; Guerrero, Mexico; and Bungo Channel and Boso Peninsula, Japan, and comparing them with SSE cyclicity. They found that 14 out of 20 SSEs occurred just prior to, and three occurred synchronous with, induced stress changes attaining their maxima. They concluded that although these stress perturbations were very small (between 0.2 and 1 kPa), the long period (14.5 months) of the modulating stress was critical in reducing the triggering threshold and inducing slow slip. In addition, they suggest that the perturbations trigger SSEs rather than earthquakes because the triggering threshold is lowered in the region of the fault plane with frictional conditions that favor SSEs.

Lowry (2006) hypothesized that periodic slow slip is a resonant fault response to stress perturbations caused by the Chandler wobble and/or other climatic loading phenomena. However, he noted that on average, hydrologic stress perturbations are at least a factor of two greater than those generated by the Chandler wobble. He noted that for slip to be triggered by hydrologic or other climatic forcing, it must respond preferentially to longer-period stress oscillations. Perfettini and Schmittbuhl (2001) found that slip resulting from an applied stress was amplified provided that the frictional conditions were near the stability transition and the period of loading was close to a critical value. Lowry (2006) called this critical period for slip amplification the natural resonant period of the fault, and showed that it depended on frictional parameters of the fault plane and the plate convergence velocity. He speculated that environmental loading by the ocean, atmosphere, and tides provided a broad enough spectrum of forcing signals to potentially match the resonant periods of all fault planes hosting SSEs.

4.17.8 Conclusions

The increasing density of continuous GPS and other geodetic observations worldwide has revealed EAS to be a ubiquitous process that constitutes a fundamental mode of strain release at

plate boundaries. For a long time, the aseismic component of plate motions had been assumed to take place primarily as steady-state motion; however, evidence is accumulating to support a model in which a large fraction of aseismic slip at plate boundaries occurs episodically rather than continuously. These episodic aseismic motions appear to have a significant influence on other, probably interrelated, tectonic processes, such as the generation of seismic tremor, pore fluid pressurization and expulsion at the seafloor (Brown et al., 2005; Davis et al., 2006), and earthquake nucleation. It is too early in the study of aseismic strain transients to make overarching conclusions about their behavior, but continued study of this phenomenon may be critical to attaining a meaningful understanding of seismogenic zones. Progress is also being made in developing and applying more geodetic techniques capable of detecting and studying EAS at plate boundaries. For example, Lambert et al. (2006) demonstrated that terrestrial absolute gravity measurements could be used in conjunction with space-based measurements to detect short-term gravity variations associated with SSEs. APGs deployed on the ocean bottom offshore the Tohoku margin detected two SSEs (in December 2008 and February 2011, a few weeks prior to the Great Tohoku earthquake and tsunami) that involved 2–3 cm of vertical deformation over 5–7 days (Ito et al., 2013). The Tohoku example demonstrates the utility of APGs to detect SSE deformation above offshore subduction thrusts.

The ongoing and intense efforts to establish and upgrade monitoring capabilities at active plate boundaries will doubtless contribute to the fast-growing database of EAS events, resulting in new discoveries and a deeper understanding of this important mode of strain release.

Acknowledgments

I thank Julie Rokosky for her collaboration and important contributions to the early EAS research efforts at UCSC. I am also grateful to Kazushige Obara, Wendy McCausland, and Chloe Peterson for providing tremor waveforms from Japan, Cascadia, and Alaska; Tim Dixon, Emily Brodsky, Thorne Lay, and Greg Beroza for the enlightening conversations about episodic aseismic slip; and John Vidale for his comments on this manuscript. NSF supported this work through awards EAR-0506463 and OCE-0841061.

References

Aguiar AC, Melbourne TI, and Scrivner CW (2009) Moment release rate of Cascadia tremor constrained by GPS. *Journal of Geophysical Research* 114: B00A05. http://dx.doi.org/10.1029/2008JB005909.

Bartlow N, Miyazaki S, Bradley A, and Segall P (2011) Space–time correlation of slip and tremor during the 2009 Cascadia slow slip event. *Geophysical Research Letters* 38. http://dx.doi.org/10.1029/2011GL048714.

Beroza GC and Ide S (2011) Slow earthquakes and nonvolcanic tremor. *Annual Review of Earth and Planetary Sciences* 39: 271–296.

Beroza GC and Jordan T (1990) Searching for slow and silent earthquakes. *Journal of Geophysical Research* 95: 2485–2510.

Bilek SL, Lay T, and Ruff LJ (2004) Radiated seismic energy and earthquake source duration variations from teleseismic source time functions for shallow subduction zone thrust earthquakes. *Journal of Geophysical Research* 109: B09308. http://dx.doi.org/10.1029/2004JB003039.

Bostock MG, Royer AA, Hearn EH, and Peacock SM (2012) Low frequency earthquakes below southern Vancouver Island. *Geochemistry, Geophysics, Geosystems* 13: Q11007.

Boyarko D and Brudzinski M (2010) Spatial and temporal patterns of nonvolcanic tremor along the southern Cascadia subduction zone. *Journal of Geophysical Research* 115: B00A22. http://dx.doi.org/10.1029/2008JB006064.

Brooks BA, Foster JH, Bevis M, Frazer LN, Wolfe CJ, and Behn M (2006) Periodic slow earthquakes on the flank of Kilauea volcano, Hawaii. *Earth and Planetary Science Letters* 246: 207–216.

Brown JR, Beroza GC, Ide S, et al. (2009) Deep low-frequency earthquakes in tremor localize to the plate interface in multiple subduction zones. *Geophysical Research Letters* 36: L19306. http://dx.doi.org/10.1029/2009GL040027.

Brown KM, Tryon MD, DeShon HR, Dorman LM, and Schwartz SY (2005) Correlated transient fluid pulsing and seismic tremor in the Costa Rica subduction zone. *Earth and Planetary Science Letters* 238: 189–203.

Brudzinski MR and Allen R (2007) Segmentation in episodic tremor and slip all along Cascadia. *Geology* 35: 907–910.

Brudzinski M, Cabral-Cano E, Correa-Mora F, DeMets C, and Marquez-Azua B (2007) Slow slip transients along the Oaxaca subduction segment from 1993 to 2007. *Geophysical Journal International* 171(2): 523–538.

Brudzinski M, Hinojosa-Prieto H, Schlanser K, et al. (2010) Nonvolcanic tremor along the Oaxaca segment of the Middle America subduction zone. *Journal of Geophysical Research: Solid Earth* 115: B00A23. http://dx.doi.org/10.1029/2008JB006061.

Bucknam RC, Plafker G, and Sharp RV (1978) Fault movement (afterslip) following the Guatemala earthquake of February 4, 1976. *Geology* 6: 170–173.

Bürgmann R, Kogan MG, Levin VE, Scholz CH, King RW, and Steblov GM (2001) Rapid aseismic moment release following the 5 December 1997 Kronotsky Kamchatka earthquake. *Geophysical Research Letters* 28: 1331–1334.

Bürgmann R, Schmidt D, Nadeau RM, et al. (2000) Earthquake potential along the northern Hayward fault, California. *Science* 289(5482): 1178–1182.

Cervelli P, Segall P, Johnson K, Lisowski M, and Miklius A (2002) Sudden aseismic fault slip on the south flank of Kilauea volcano. *Nature* 415: 1014–1018.

Cifuentes IL and Silver PG (1989) The 1960 Chilean earthquake. *Journal of Geophysical Research* 94: 665–680.

Cohen SC and Freymueller JT (2004) Crustal deformation in southcentral Alaska: The 1964 Prince William Sound earthquake subduction zone. *Advances in Geophysics* 47: 1–63.

Correa-Mora F, DeMets C, Cabral-Cano E, Diaz-Molina O, and Marquez-Azua B (2009) Transient deformation in southern Mexico in 2006 and 2007: Evidence for distinct deep-slip patches beneath Guerrero and Oaxaca. *Geochemistry, Geophysics, Geosystems* 10: 1–12.

Correa-Mora F, DeMets C, Cabral-cano E, Marquez-Azua B, and Diaz-Molina O (2008) Interplate coupling and transient slip along the subduction interface beneath Oaxaca, Mexico. *Geophysical Journal International* 175: 269–290. http://dx.doi.org/10.1111/j.1365-246X.2008.03910.x.

Davis EE, Becker K, Wang K, Obara K, Ito Y, and Kinoshita M (2006) A discrete episode of seismic and aseismic deformation of the Nankai trough subduction zone accretionary prism and incoming Philippine Sea plate. *Earth and Planetary Science Letters* 242: 73–84.

Delahaye E, Townend J, Reyners M, and Rogers G (2009) Microseismicity but no tremor accompanying slow slip in the Hikurangi subduction zone, New Zealand. *Earth and Planetary Science Letters* 277: 21–28.

Dieterich JH (1979) Modeling of rock friction 1. Experimental results and constitutive equations. *Journal of Geophysical Research* 84: 2161–2168.

Douglas A, Beavan J, Wallace L, and Townend J (2005) Slow slip on the northern Hikurangi subduction interface, New Zealand. *Geophysical Research Letters* 32. http://dx.doi.org/10.1029/2005GL023607.

Dragert H, Hyndman RD, Rogers GC, and Wang K (1994) Current deformation and the width of the seismogenic zone of the northern Cascadia subduction thrust. *Journal of Geophysical Research* 99(B1): 653–658.

Dragert H, Wang K, and James TS (2001) A silent slip event on the deeper Cascadia subduction interface. *Science* 292: 1525–1528.

Dragert H, Wang K, and Rogers G (2004) Geodetic and seismic signatures of episodic tremor and slip in the northern Cascadia subduction zone. *Earth, Planets and Space* 56: 1143–1150.

Evans KF, Burford RO, and King GCP (1981) Propagating episodic creep and aseismic slip behavior of the Calaveras Fault north of Hollister, California. *Journal of Geophysical Research* 86: 3721–3735.

Forsyth DW, Yang Y, Mangriotis MD, and Shen Y (2003) Coupled seismic slip on adjacent oceanic transform faults. *Journal of Geophysical Research* 30. http://dx.doi.org/10.1029/2002GL016454.

Galehouse JS (2002) Data from theodolite measurements of creep rates on San Francisco Bay Region faults, California: 1979–2001. *U. S. Geological Survey Open-File Report, 02–225*.

Galehouse JS and Lienkaemper JJ (2003) Inferences drawn from two decades of alignment array measurements of creep on faults in the San Francisco Bay region. *Bulletin of the Seismological Society of America* 93: 2415–2433.
Gao SS, Silver PG, and Linde AT (2000) Analysis of deformation data at Parkfield, California: Detection of a long-term strain transient. *Journal of Geophysical Research* 105: 2955–2968.
Ghosh A, Vidale JE, Sweet JR, Creager KC, and Wech AG (2009) Tremor patches in Cascadia revealed by seismic array analysis. *Geophysical Research Letters* 36: L17316. http://dx.doi.org/10.1029/2009GL039080.
Ghosh A, Vidale JE, Sweet JR, Creager K, Wech A, and Houston H (2010) Tremor bands sweep Cascadia. *Geophysical Research Letters* 37: L08301. http://dx.doi.org/10.1029/2009GL042301.
Gladwin MT (2004) *Maintenance, Data Archive and Analysis of Existing Low-frequency GTSM Installations in California, National Earthquake Hazard Reduction Program Annual Project Summary*. Reston, VA: The U.S. Geological Survey.
Gladwin MT, Gwyther RL, Hart RHG, and Breckenridge K (1994) Measurements of the strain field associated with episodic creep events on the San Andreas Fault near San Juan Bautista, California. *Journal of Geophysical Research* 99: 4559–4565.
Graham SE, Cabral-Cano E, Kostoglodov V, et al. (2012) Animations and modeling of the 2012 slow slip event beneath southern Mexico and the accompanying 20 March 2012 $M_w = 7.4$ Ometepec earthquake. In: Abstract S33B-2528 Presented at 2012 Fall Meeting, San Francisco, CA: AGU.
Gwyther RL, Gladwin MT, Mee M, and Hart RHG (1996) Anomalous tensor strain at Parkfield during 1993–1994. *Geophysical Research Letters* 23: 2425–2428.
Gwyther RL, Thurber CH, Gladwin MT, and Mee M (2000) Seismic and aseismic observations of the 12th August 1998 San Juan Bautista, California M5.3 earthquake. In: *Paper Presented at 3rd San Andreas Fault Conference, Stanford University, Stanford, CA*.
Hamilton S and Shennan I (2005) Late Holocene great earthquakes and relative sea level change at Kenai, southern Alaska. *Journal of Quaternary Science* 20: 95–111.
Heki K and Kataoka T (2008) On the biannually repeating slow-slip events at the Ryukyu Trench, southwestern Japan. *Journal of Geophysical Research* 113: B11402. http://dx.doi.org/10.1029/2008JB005739.
Heki K, Miyazaki S, and Tsuji H (1997) Silent fault slip following an interplate thrust earthquake at the Japan Trench. *Nature* 386: 595–598.
Hirose H, Asano Y, Obara K, et al. (2010) Slow earthquakes linked along dip in the Nankai subduction zone. *Science* 330: 1502.
Hirose H, Hirahara K, Kimata F, Fujii N, and Miyazaki S (1999) A slow thrust slip event following the two 1996 Hyuganada earthquakes beneath the Bungo Channel, southwest Japan. *Geophysical Research Letters* 26(21): 3237–3240. http://dx.doi.org/10.1029/1999GL010999.
Hirose H and Obara K (2005) Repeating short- and long-term slow slip events with deep tremor activity around the Bungo channel region, southwest Japan. *Earth, Planets and Space* 57: 961–972.
Hirose H and Obara K (2006) Short-term slow slip and correlated tremor episodes in the Tokai region, central Japan. *Geophysical Research Letters* 33: L17311. http://dx.doi.org/10.1029/2006GL026579.
Hsu Y-J, Simons M, Avouac J-P, et al. (2006) Frictional afterslip following the 2005 Nias-Simeulue earthquake, Sumatra. *Science* 312: 1921–1926.
Hutton W, DeMets C, Sanchez O, Suarez G, and Stock J (2001) Slip kinematics and dynamics during and after the 1995 October 9 $M_w = 8.0$ Colima-Jalisco earthquake, Mexico, from GPS geodetic constraints. *Geophysical Journal International* 146: 637–658.
Hyndman RD, Wang K, and Yamano M (1995) Thermal constraints on the seismogenic portion of the southwestern Japan subduction thrust. *Journal of Geophysical Research* 100: 15373–15392.
Ide S (2012) Variety and spatial heterogeneity of tectonic tremor worldwide. *Journal of Geophysical Research* 117: B03302. http://dx.doi.org/10.1029/2011JB008840.
Igarashi T, Matsuzawa T, and Hasegawa A (2003) Repeating earthquakes and interplate aseismic slip in the northeastern Japan subduction zone. *Journal of Geophysical Research* 108: 2249. http://dx.doi.org/10.1029/2002JB001920.
Ito T and Hashimoto M (2004) Spatiotemporal distribution of interplate coupling in southwest Japan from inversion of geodetic data. *Journal of Geophysical Research* 109: B02315. http://dx.doi.org/10.1029/2002JB002358.
Ito T, Yoshioka S, and Miyazaki S (1999) Interplate coupling in southwest Japan deduced form inversion analysis of GPS data. *Physics of the Earth and Planetary Interiors* 115: 17–34.
Ito Y, et al. (2013) Episodic slow slip events in the Japan subduction zone before the 2011 Tohoku-Oki earthquake. *Tectonophysics* 600: 14–26.
Jiang Y (2012) *Detection of Nonlinear Crustal Movements Using Global Positioning System*. PhD Thesis, University of Miami, Open Access Dissertations, Paper 844.
Jiang Y, Wdowinski S, Dixon TH, Hackl M, Protti M, and Gonzalez V (2012) Slow slip events in Costa Rica detected by continuous GPS observations, 2002–2011. *Geochemistry, Geophysics, Geosystems* 13: Q04006. http://dx.doi.org/10.1029/2012GC004058.
Johnston MJS (1997) Near-silent earthquakes on the San Andreas Fault system. Fall Meeting Supplement 156, *Eos, Transactions American Geophysical Union* 78(46).
Johnston MJS, Gwyther RL, Linde AT, Gladwin MT, Myren GD, and Mueller RJ (1996) Another slow earthquake on the San Andreas Fault triggered by a M4.7 earthquake on April 19, 1996. Fall Meeting Supplement 515, *Eos, Transactions American Geophysical Union* 77(46).
Johnston M and Linde A (2002) Implications of crustal strain during conventional, slow, and silent earthquakes. In: Lee WHK, et al. (eds.) *International Handbook of Earthquake and Engineering Seismology, Part A*. *International Geophysical Services*, vol. 81, pp. 589–605. New York: Elsevier.
Kanamori H and Cipar J (1974) Focal process of the great Chilean earthquake May 22, 1960. *Physics of the Earth and Planetary Interiors* 9: 127–136.
Kanamori H and Kikuchi M (1993) The 1992 Nicaragua earthquake: A slow tsunami earthquake associated with subducted sediments. *Nature* 361(6414): 714–716.
Kanamori H and Stewart GS (1976) Mode of strain release along the Gibbs fracture zone, mid-Atlantic ridge. *Physics of the Earth and Planetary Interiors* 11: 312–332.
Kao H, Shan S, Dragert H, Rogers G, Cassidy JF, and Ramachandran K (2005) A wide depth distribution of seismic tremors along the northern Cascadia margin. *Nature* 436: 841–844.
Kato A, Obara K, Igarashi T, Tsuruoka H, Nakagawa S, and Hirata N (2012) Propagation of slow slip leading up to the 2011 M_w 9.0 Tohoku-Oki earthquake. *Science* 335: 705–708. http://dx.doi.org/10.1126/science.1215141.
Kawasaki I, Asai Y, and Tamura Y (2001) Space–time distribution of interplate moment release including slow earthquakes and the seismogeodetic coupling in the Sanriku-Oki region along the Japan trench. *Tectonophysics* 330: 267–283.
Kawasaki I, Asai Y, Tamura Y, et al. (1995) The 1992 Sanriku-Oki, Japan, ultra-slow earthquake. *Journal of Physics of the Earth* 43: 105–116.
Kim M, Schwartz S, and Bannister S (2011) Non-volcanic tremor associated with the March 2010 Gisborne slow slip event at the Hikurangi subduction margin, New Zealand. *Geophysical Research Letters* 38: L14301. http://dx.doi.org/10.1029/2011GL048400.
Kostoglodov V, Singh SK, Santiago JA, et al. (2003) A large silent earthquake in the Guerrero seismic gap, Mexico. *Geophysical Research Letters* 30: 1807. http://dx.doi.org/10.1029/2003GL017219.
Kreemer C, Blewitt G, and Maerten F (2006) Co- and postseismic deformation of the 28 March 2005 Nias Mw 8.7 earthquake from continuous GPS data. *Geophysical Research Letters* 33: L07307.
Kurashimo E and Naoshi H (2004) Low Vp and Vp/Vs zone beneath the northern Fossa Magna basin, central Japan, derived from a dense array observation. *Earth, Planets and Space* 56(12): 1301–1308.
Kuroki H, Ito HM, Takayama H, and Yoshida A (2004) 3-D simulation of the occurrence of slow slip events in the Tokai region with a rate- and state-dependent friction law. *Bulletin of the Seismological Society of America* 94: 2037–2050.
LaBonte A, Brown K, and Fialko Y (2009) Hydrologic detection and finite element modeling of a slow slip event in the Costa Rica prism toe. *Journal of Geophysical Research: Solid Earth* 114. http://dx.doi.org/10.1029/2008JB005806.
Lambert A, Courtier N, and James TS (2006) Long-term monitoring by absolute gravimetry: Tides to postglacial rebound. *Journal of Geodynamics* 41: 307–317.
Langbein J, Gwyther RL, Hart RHG, and Gladwin MT (1999) Slip-rate increase at Parkfield in 1993 detected by high-precision EDM and borehole tensor strainmeters. *Geophysical Research Letters* 26: 2529–2532.
Larson KM, Kostoglodov V, Lowry A, et al. (2004) Crustal deformation measurements in Guerrero, Mexico. *Journal of Geophysical Research* 109: B04409. http://dx.doi.org/10.1029/2003JB002843.
Larson K, Kostoglodov V, Miyazaki S, and Santiago J (2007) The 2006 aseismic slow slip event in Guerrero, Mexico: New results from GPS. *Geophysical Research Letters* 34(13).
Linde AT, Gladwin M, Johnston M, Gwyther R, and Bilham R (1996) A slow earthquake sequence on the San Andreas fault. *Nature* 383: 65–68.
Linde AT and Sacks SI (2002) Slow earthquakes and great earthquakes along the Nankai trough. *Earth and Planetary Science Letters* 203: 265–275.
Linde AT and Silver PG (1989) Elevation changes and the great 1960 Chilean earthquake: Support for aseismic slip. *Geophysical Research Letters* 16: 1305–1308.
Linde AT, Suyehiro K, Satoshi M, Sacks SI, and Takagi A (1988) Episodic aseismic earthquake precursors. *Nature* 334(6182): 513–515.
Liu Y and Rice JR (2005) Aseismic slip transients emerge spontaneously in three-dimensional rate and state modeling of subduction earthquake sequences. *Journal of Geophysical Research* 110: B08307. http://dx.doi.org/10.1029/2004JB003424.
Liu Y and Rice JR (2007) Spontaneous and triggered aseismic deformation transients in a subduction fault model. *Journal of Geophysical Research* 112: B09404.
Lowry AR (2006) Resonant slow fault slip in subduction zones forced by climatic load stress. *Nature* 442(7104): 802–805.

Lowry AR, Larson KM, Kostoglodov V, and Bilham R (2001) Transient fault slip in Guerrero, southern Mexico. *Geophysical Research Letters* 28: 3753–3756.

Marone CJ, Scholtz CH, and Bilham R (1991) On the mechanics of afterslip. *Journal of Geophysical Research* 96: 8441–8452.

Matsubara M, Yagi Y, and Obara K (2005) Plate boundary slip associated with the 2003 Off-Tokachi earthquake based on small repeating earthquakes. *Geophysical Research Letters* 32: L08316. http://dx.doi.org/10.1029/2004GL022310.

Matsumura S (2006) Seismic activity changes progressing simultaneously with slow-slip in the Tokai area. *Tectonophysics* 417(1–2): 5–15.

McCausland W, Malone S, and Johnson D (2005) Temporal and spatial occurrence of deep non-volcanic tremor: From Washington to northern California. *Geophysical Research Letters* 32(24): L24311. http://dx.doi.org/10.1029/2005GL024349.

McGuire JJ, Boettcher MS, and Jordan TH (2005) Foreshock sequences and short-term earthquake predictability on East Pacific Rise transform faults. *Nature* 434: 457–461.

McGuire JJ, Ihmle PF, and Jordan TH (1996) Time-domain observations of a slow precursor to the 1994 Romanche transform earthquake. *Science* 274: 82–85.

McGuire JJ and Jordan TH (2000) Further evidence for the compound nature of slow earthquakes: The Prince Edward Island earthquake of April 28, 1997. *Journal of Geophysical Research* 105: 7819–7828.

McGuire JJ and Segall P (2003) Imaging of aseismic fault slip transients recorded by dense geodetic networks. *Geophysical Journal International* 155: 778–788.

Melbourne TI, Szeliga WM, Miller MM, and Santillan VM (2005) Extent and duration of the 2003 Cascadia slow earthquake. *Geophysical Research Letters* 32(4): L04301. http://dx.doi.org/10.1029/2004GL021790.

Melbourne TI and Webb FH (2002) Precursory transient slip during the 2001 $M_w = 8.4$ Peru earthquake sequence from continuous GPS. *Geophysical Research Letters* 29(21): 2032. http://dx.doi.org/10.1029/2002GL015533.

Melbourne TI, Webb FH, Stock JM, and Riegber C (2002) Rapid postseismic transients in subduction zones from continuous GPS. *Journal of Geophysical Research* 107: B102241. http://dx.doi.org/10.1029/2001JB000555.

Miller MM, Melbourne T, Johnson DJ, and Sumner WQ (2002) Periodic slow earthquakes from the Cascadia subduction zone. *Science* 295: 2423.

Miyazaki S, McGuire JJ, and Segall P (2003) A transient subduction zone slip episode in southwest Japan observed by the nationwide GPS array. *Journal of Geophysical Research* 108. http://dx.doi.org/10.1029/2001JB000456.

Miyazaki S, Segall P, McGuire JJ, Kato T, and Hatanaka Y (2006) Spatial and temporal evolution of stress and slip rate during the 2000 Tokai slow earthquake. *Journal of Geophysical Research* 111: B03409. http://dx.doi.org/10.1029/2004JB003426.

Montgomery-Brown E, Segall P, and Miklius A (2009) Kilauea slow slip events: Identification, source inversions, and relation to seismicity. *Journal of Geophysical Research: Solid Earth* 114. http://dx.doi.org/10.1029/2008JB006074.

Murray JR and Segall P (2005) Spatiotemporal evolution of a transient slip event on the San Andreas fault near Parkfield, California. *Journal of Geophysical Research* 110. http://dx.doi.org/10.1029/2005JB003651.

Nadeau R and Dolenc D (2005) Nonvolcanic tremors deep beneath the San Andreas Fault. *Science* 307: 389.

Nadeau R and Johnson LR (1998) Seismological studies at Parkfield VI: Moment release rates and estimates of source parameters for small repeating earthquakes. *Bulletin of the Seismological Society of America* 88: 790–814.

Nadeau RM and McEvilly TV (1999) Fault slip rates at depth from recurrence intervals of repeating microearthquakes. *Science* 285: 718–721.

Nadeau RM and McEvilly TV (2004) Periodic pulsing of characteristic microearthquakes on the San Andreas Fault. *Science* 303: 220–222.

Newman AV and Okal EA (1998) Teleseismic estimates of radiated seismic energy: The E/M_0 discriminant for tsunami earthquakes. *Journal of Geophysical Research* 103(11): 26885–26898.

Norabuena E, Dixon TH, Schwartz SY, et al. (2004) Geodetic and seismic constraints on some seismogenic zone processes in Costa Rica. *Journal of Geophysical Research* 109: B11403. http://dx.doi.org/10.1029/2003JB002931.

Obara K (2002) Nonvolcanic deep tremor associated with subduction in southwest Japan. *Science* 296: 1679–1681.

Obara K (2010) Phenomenology of deep slow earthquake family in southwest Japan: Spatiotemporal characteristics and segmentation. *Journal of Geophysical Research* 115: B00A25. http://dx.doi.org/10.1029/2008JB006048.

Obara K (2011) Characteristics and interactions between non-volcanic tremor and related slow earthquakes in the Nankai subduction zone, southwest Japan. *Journal of Geodynamics* 52(3–4): 229–248.

Obara K and Hirose H (2006) Non-volcanic deep low-frequency tremors accompanying slow slips in the southwest Japan subduction zone. *Tectonophysics* 417(1–2): 33–51.

Obara K, Hirose H, Yamamizu F, and Kasahara K (2004) Episodic slow slip events accompanied by non-volcanic tremors in southwest Japan subduction zone. *Geophysical Research Letters* 31. http://dx.doi.org/10.1029/2004GL020848.

Ohta Y, Freymueller JT, Hreinsdóttir S, and Suito H (2006) A large slow slip event and the depth of the seismogenic zone in the south central Alaska subduction zone. *Earth and Planetary Science Letters* 247(1–2): 108–116.

Ohta Y, Kimata F, and Sagiya T (2004) Reexamination of the interplate coupling in the Tokai region, central Japan, based on the GPS data in 1997–2002. *Geophysical Research Letters* 31: L24604. http://dx.doi.org/10.1029/2004GL021404.

Okada Y (1985) Surface deformation due to shear and tensile faults in a half-space. *Bulletin of the Seismological Society of America* 75: 1135–1154.

Outerbridge KC, Dixon TH, Schwartz SY, et al. (2010) A tremor and slip event on the Cocos-Caribbean Subduction zone as measured by a GPS and Seismic Network on the Nicoya Peninsula, Costa Rica. *Journal of Geophysical Research* http://dx.doi.org/10.1029/2009JB006845.

Ozawa S, Hatanaka Y, Kaidzu M, Murakami M, Imakiire T, and Ishigaki Y (2004) Aseismic slip and low-frequency earthquakes in the Bungo Channel, southwestern Japan. *Geophysical Research Letters* 31: 7609. http://dx.doi.org/10.1029/2003GL016665.

Ozawa S, Miyazaki S, Hatanaka Y, Imakiire T, Kaidzu M, and Murakami M (2003) Characteristic silent earthquakes in the eastern part of the Boso peninsula, Central Japan. *Geophysical Research Letters* 30: 1283. http://dx.doi.org/10.1029/2002GL019381.

Ozawa S, Murakami M, Kaidzu M, et al. (2002) Detection and monitoring of ongoing aseismic slip in the Tokai region, Central Japan. *Science* 298: 1009–1012.

Ozawa S, Nishimura T, Munekane H, et al. (2012) Preceding, coseismic, and postseismic slips of the 2011 Tohoku earthquake, Japan. *Journal of Geophysical Research: Solid Earth* 117: B07404. http://dx.doi.org/10.1029/2011JB009120.

Ozawa S, Suito H, and Tobita M (2007) Occurrence of quasi-periodic slow-slip off the east coast of the Boso peninsula, Central Japan. *Earth, Planets and Space* 59(12): 1241–1245.

Payero J, Kostoglodov V, Shapiro N, et al. (2008) Nonvolcanic tremor observed in the Mexican subduction zone. *Geophysical Research Letters* 35(7).

Pelayo AM and Wiens DA (1992) Tsunami earthquakes: Slow thrust-faulting events in the accretionary wedge. *Journal of Geophysical Research* 97(11): 15321–15337.

Perez-Campos X, McGuire JJ, and Beroza GC (2003) Resolution of the slow earthquake/apparent stress paradox for oceanic transform fault earthquakes. *Journal of Geophysical Research* 108. http://dx.doi.org/10.1029/2002JB002312.

Perfettini H, Avouac JP, Tavera H, et al. (2010) Seismic and aseismic slip on the central Peru megathrust. *Nature* 465. http://dx.doi.org/10.1038/nature09062.

Perfettini H and Schmittbuhl J (2001) Periodic loading on a creeping fault: Implications for tides. *Geophysical Research Letters* 28(3): 435–438.

Peterson CL and Christensen DH (2009) Possible relationship between nonvolcanic tremor and the 1998–2001 slow slip event, south central Alaska. *Journal of Geophysical Research* 114: B06302. http://dx.doi.org/10.1029/2008JB006096.

Pollitz FF and Johnston MJS (2006) Direct test of static stress versus dynamic stress triggering of aftershocks. *Geophysical Research Letters* 33. http://dx.doi.org/10.1029/2006GL026764.

Pritchard ME and Simons M (2006) An aseismic slip pulse in northern Chile and along-strike variations in seismogenic behavior. *Journal of Geophysical Research* 111. http://dx.doi.org/10.1029/2006JB004258.

Protti M, Gonzales V, Kato T, et al. (2004) A creep event on the shallow interface of the Nicoya Peninsula, Costa Rica seismogenic zone. Fall Meeting Supplement, *EOS, Transactions, American Geophysical Union* 85: Abstract S441D-07.

Radiguet M, Cotton F, Vergnolle M, et al. (2011) Spatial and temporal evolution of a long term slow slip event: The 2006 Guerrero Slow Slip Event. *Geophysical Journal International* 184(2): 816–828.

Roeloffs E (2001) Creep rate changes at Parkfield, California, 1966–1999: Seasonal, precipitation induced, and tectonic. *Journal of Geophysical Research* 106: 16525–16547.

Roeloffs EA (2006) Evidence for aseismic deformation rate changes prior to earthquakes. *Annual Review of Earth and Planetary Sciences* 34: 591–627.

Rogers G and Dragert H (2003) Episodic tremor and slip on the Cascadia subduction zone: The chatter of silent slip. *Science* 300: 1942–1943.

Rubin AM (2008) Episodic slow slip events and rate-and-state friction. *Journal of Geophysical Research* 113: B11414.

Rubinstein JL, Shelly DR, and Ellsworth WL (2010) Non-volcanic tremor: A window into the roots of fault zones. In: Cloetingh S and Negendank J (eds.) *New Frontiers in Integrated Solid Earth Sciences*, pp. 287–314. Dordrecht, Netherlands: Springer. http://dx.doi.org/10.1007/978-90-481-2737-5_8.

Sacks IS, Linde AT, Snoke JA, and Suyehiro S (1981) A slow earthquake sequence following the Izu-Oshima earthquake of 1978. In: Simpson DW and Richards PG (eds.) *Earthquake Prediction: An International Review*. *Maurice Ewing Series*, vol. 4. Washington, D.C.: AGU.

Sacks IS, Suyhiro S, Evertson DW, and Yamagishi Y (1971) Sacks–Evertson strainmeter: Its installation in Japan and some preliminary results concerning strainsteps. *Papers in Meteorology and Geophysics* 22: 195–208.
Sagiya T (2004) Interplate coupling in the Kanto District, central Japan, and the Boso Peninsula silent earthquake in May 1996. *Pure and Applied Geophysics* 161: 2327–2342.
Sagiya T, Yoshimura A, Iwata E, et al. (1995) Establishment of permanent GPS observation network and crustal deformation monitoring in the southern Kanto and Tokai areas. *Bulletin of the Geographical Survey Institute* 41: 105–118.
Satake K, Wang K, and Atwater BF (2003) Fault slip and seismic moment of the 1700 Cascadia earthquake inferred from Japanese tsunami descriptions. *Journal of Geophysical Research* 108(B11): 2535. http://dx.doi.org/10.1029/2003JB002521.
Sato T, Imanishi K, Kato N, and Sagiya T (2004) Detection of a slow slip event from small signal in GPS data. *Geophysical Research Letters* 31: L05606. http://dx.doi.org/10.1029/2004GL019514.
Schmidt DA, Burgmann R, Nadeau RM, and d'Alessio M (2005) Distribution of aseismic slip rate on the Hayward fault inferred from seismic and geodetic data. *Journal of Geophysical Research* 110. http://dx.doi.org/10.1029/2004JB003397.
Schwartz SY and Rokosky JM (2007) Slow slip events and seismic tremor at circum-Pacific subduction zones. *Reviews of Geophysics* 45: RG3004.
Segall P, Desmarais EK, Shelly D, Miklius A, and Cervelli P (2006) Earthquakes triggered by silent slip events on Kilauea Volcano, Hawaii. *Nature* 442(7098): 71–74.
Segall P, Rubin AM, Bradley AM, and Rice JR (2010) Dilatant strengthening as a mechanism for slow slip events. *Journal of Geophysical Research* 115: B12305.
Sekine S, Hirose H, and Obara K (2010) Short-term slow slip events correlated with non-volcanic tremor episodes in southwest Japan. *Journal of Geophysical Research* http://dx.doi.org/10.1029/2008JB006059.
Shelly DR (2009) Possible deep fault slip preceding the 2004 Parkfield earthquake, inferred from detailed observations of tectonic tremor. *Geophysical Research Letters* 36. http://dx.doi.org/10.1029/2009gl039589.
Shelly DR (2010) Migrating tremors illuminate complex deformation beneath the seismogenic San Andreas Fault. *Nature* 463. http://dx.doi.org/10.1038/nature08755.
Shelly DR, Beroza GC, and Ide S (2007) Non-volcanic tremor and low-frequency earthquake swarms. *Nature* 446(7133): 305–307.
Shelly DR, Beroza GC, Satoshi I, and Nakamula S (2006) Low-frequency earthquakes in Shikoku, Japan and their relationship to episodic tremor and slip. *Nature* 442(7099): 188–191.
Shelly DR and Hardebeck JL (2010) Precise tremor source locations and amplitude variations along the lower-crustal central San Andreas Fault. *Geophysical Research Letters* 37: L14301. http://dx.doi.org/10.1029/2010GL043672.
Shen Z, Wang Q, Burgmann R, Wan Y, and Ning J (2005) Pole-tide modulation of slow slip events at circum-Pacific subduction zones. *Bulletin of the Seismological Society of America* 95(5): 2009–2015.
Shennan I, Long AJ, Rutherford MM, Innes JB, Green FM, and Walker KJ (1998) Tidal marsh stratigraphy, sea-level change and large earthquakes II: Submergence events during the last 3500 years at Netarts Bay, Oregon, USA. *Quaternary Science Reviews* 17: 365–394.
Shibazaki B and Iio Y (2003) On the physical mechanism of silent slip events along the deeper part of the seismogenic zone. *Geophysical Research Letters* 30(9): 1489. http://dx.doi.org/10.1029/2003GL017047.
Smith SW and Wyss M (1968) Displacement on the San Andreas Fault subsequent to the 1966 Parkfield earthquake. *Bulletin of the Seismological Society of America* 58: 1955–1973.
Song T, Helmberger D, Brudzinski M, et al. (2009) Subducting slab ultra-slow velocity layer coincident with silent earthquakes in southern Mexico. *Science* 324: 502–506.
Subarya C, Chlieh M, Prawirodirdjo L, et al. (2006) Plate-boundary deformation associated with the great Sumatra–Andaman earthquake. *Nature* 440: 46–51.
Suito H and Ozawa T (2009) Transient crustal deformation in the Tokai district. *Journal of the Seismological Society of Japan* 2(61): 113–135, in Japanese with English abstract.
Suwa Y, Miura S, Hasegawa A, Sato T, and Tachibana K (2006) Interplate coupling beneath NE Japan inferred from three-dimensional displacement field. *Journal of Geophysical Research* 111: B04402. http://dx.doi.org/10.1029/2004JB003203.
Sylvester AG (1993) Investigation of nearfield postseismic slip following the Mw 7.3 Landers earthquake sequence of 28 June 1992, California. *Geophysical Research Letters* 20(11): 1079–1082.
Szeliga W, Melbourne T, Miller M, and Santillian V (2004) Southern Cascadia episodic slow earthquakes. *Geophysical Research Letters* 31: L16602. http://dx.doi.org/10.1029/2004GL020824.
Uchida N, Matsuzawa T, Igarashi T, and Hasagawa A (2003) Interplate quasi-static slip off Sanriku, NE Japan, estimated from repeating earthquakes. *Geophysical Research Letters* 30. http://dx.doi.org/10.1029/2003GL017452.
Ueda H, Ohtake M, and Sato H (2001) Afterslip of the plate interface following the 1978 Miyagi-Oki Japan earthquake, as revealed from geodetic measurement data. *Tectonophysics* 338: 45–57.
Uhrhammer R, Gee LS, Murray MH, Dreger D, and Romanowicz B (1999) The Mw 5.1 San Juan Bautista, California earthquake of 12 August 1998. *Seismological Research Letters* 70(1): 10–18.
Vergnolle M, Walpersdorf A, Kostoglodov V, et al. (2010) Slow slip events in Mexico revised from the processing of 11 year GPS observations. *Journal of Geophysical Research: Solid Earth* 115: 1–18.
Vidale J, Hotovec A, Ghosh A, Creager K, and Gomberg J (2011) Tiny intraplate earthquakes triggered by nearby episodic tremor and slip in Cascadia. *Geochemistry, Geophysics, Geosystems* 12. http://dx.doi.org/10.1029/2011GC003559.
Vidale JE and Shearer PM (2006) A survey of 71 earthquake bursts across southern California: Exploring the role of pore fluid pressure fluctuations and aseismic slip as drivers. *Journal of Geophysical Research* 111. http://dx.doi.org/10.1029/2005JB004034.
Vigny C, Simons WJF, Abu S, et al. (2006) Insight into the 2004 Sumatra–Andaman earthquake from GPS measurements in southeast Asia. *Nature* 436: 201–206.
Vigny C, Socquet A, Peyrat S, Ruegg J-C, and 26 others (2011) The 2010 Mw 8.8 Maule megathrust earthquake of Central Chile, monitored by GPS. *Science* 332: 1417. http://dx.doi.org/10.1126/science.1204132.
Wald DJ and Heaton TH (1994) Spatial and temporal distribution of slip for the 1992 Landers, California, earthquake. *Bulletin of the Seismological Society of America* 84(3): 668–691.
Wald DJ and Somerville PG (1995) Variable slip rupture model for the great 1923 Kanto, Japan earthquake: Geodetic and body waveform analysis. *Bulletin of the Seismological Society of America* 85: 159–177.
Wallace L and Beavan J (2010) Diverse slow slip behavior at the Hikurangi subduction margin, New Zealand. *Journal of Geophysical Research: Solid Earth* 115: B12402. http://dx.doi.org/10.1029/2010JB007717.
Wallace LM, Beavan J, Bannister S, and Williams C (2012) Simultaneous long-term and short-term slow slip events at the Hikurangi subduction margin, New Zealand: Implications for processes that control slow slip event occurrence, duration, and migration. *Journal of Geophysical Research* 117: B11402. http://dx.doi.org/10.1029/2012JB009489.
Walpersdorf A, Cotte N, Kostoglodov V, et al. (2011) Persistent tremor within the northern Costa Rica seismogenic zone. *Geophysical Research Letters* 38: L01307. http://dx.doi.org/10.1029/2010GL045586.
Walter JI, Schwartz SY, Protti M, and Gonzalez V (2011) Persistent tremor within the northern Costa Rica seismogenic zone. *Geophysical Research Letters* 38: L01307. http://dx.doi.org/10.1029/2010GL045586.
Wei M, McGuire J, and Richardson E (2012) A slow slip event in the south central Alaska Subduction Zone and related seismicity anomaly. *Geophysical Research Letters* 39. http://dx.doi.org/10.1029/2012GL052351.
Wesson RL (1988) Dynamics of fault creep. *Journal of Geophysical Research* 93: 8929–8951.
Williams PL and Magistrale HW (1989) Slip along the Superstition Hills fault associated with the 24 November 1987 Superstition Hills, California, earthquake. *Bulletin of the Seismological Society of America* 79: 390–410.
Yagi Y, Kikuchi M, and Nishimura T (2003) Co-seismic slip, post-seismic slip, and largest aftershock associated with the 1994 Sanriku-haruka-oki, Japan, earthquake. *Geophysical Research Letters* 30(22): 2177. http://dx.doi.org/10.1029/2003GL018189.
Yamamoto E, Matsumura S, and Ohkubo T (2005) A slow slip event in the Tokai area detected by tilt and seismic observation and its possible recurrence. *Earth, Planets and Space* 57: 917–923.
Yamanaka Y and Kikuchi M (2003) Source process of the recurrent Tokachi-oki earthquake on September 26, 2003, inferred from teleseismic body waves. *Earth, Planets and Space* 55: e1–e24.
Yamanaka Y and Kikuchi M (2004) Asperity map along the subduction zone along northeastern Japan inferred from regional seismic data. *Journal of Geophysical Research* 109(7). http://dx.doi.org/10.1029/2003JB002683.
Yoshida A, Hososno K, Tsukakoshi T, Kobayashi A, Takayama H, and Wiemer S (2006) Change in seismic activity in the Tokai region related to weakening and strengthening of the interplate coupling. *Tectonophysics* 417: 17–31.
Yoshida S and Kato N (2003) Episodic aseismic slip in a two-degree-of-freedom block-spring model. *Geophysical Research Letters* 30(13): 1681. http://dx.doi.org/10.1029/2003GL017439.

This page intentionally left blank

4.18 Global Seismicity: Results from Systematic Waveform Analyses, 1976–2012

G Ekström, Columbia University, Palisades, NY, USA

© 2015 Elsevier B.V. All rights reserved.

4.18.1	Introduction	467
4.18.2	The CMT Method	467
4.18.3	Aspects of Global Seismicity	468
4.18.4	Recent Discoveries and Future Directions	471
Acknowledgments		473
References		474

4.18.1 Introduction

Our knowledge of the geographic patterns of earthquake activity away from populated areas evolved with the global installation of seismographic stations during the first half of the twentieth century. In their classic 1941 paper, *Seismicity of the Earth*, and their 1949 book, Gutenberg and Richter (1941, 1949) documented the then-current understanding of seismic zones around the world, and some 25 years later, the spatial features of earthquake activity they described were given a comprehensive explanation in the context of plate tectonics. Integral to this explanation were the development of a kinematic model of the earthquake source in terms of fault motion and techniques of inferring the geometry of faulting from features in seismic recordings, initially from the first motions of *P* waves.

While the fundamental aspects of global seismicity are well understood, the determination of source characteristics for individual earthquakes (i.e., location, magnitude, and source mechanism) remains important. The details of the seismicity of different regions are highly variable, and neither seismicity patterns nor focal mechanisms can be predicted with much precision from a simple characterization in terms of convergent, divergent, or transcurrent plate motion. Higher-order descriptions of the deformation field are needed, and these are best constrained by comprehensive investigations of the patterns of earthquake activity and strain release.

Through the 1970s, the dominant mode of seismic recording was analog, and the determination of a single earthquake focal mechanism required a significant research effort. Access to the seismograms was the initial challenge, and any quantitative analysis required manual digitization of the seismograms. The low dynamic range of the analog recordings also generally restricted the analysis for teleseismic distances to earthquakes with magnitudes >6. The development and deployment of digital instruments with low noise in the mid-1970s provided the initial conditions for a more comprehensive and quantitative analysis of global seismicity.

Concurrent with technical progress in the high-fidelity recording of seismic waves, theoretical and computational advances in the calculation of theoretical waveforms made possible new methods of source analysis, in particular those based on the direct comparison and correlation of observed and predicted long-period waveforms. A very successful approach was that developed by Adam M. Dziewonski and his colleagues at Harvard University, who in 1981 (Dziewonski et al., 1981) published the Centroid-Moment-Tensor (CMT) algorithm for the systematic analysis of global and regional seismicity. The ongoing systematic application of the CMT method to data from the Global Seismographic Network (GSN) in the Global CMT Project (website www.globalcmt.org) has led to the accumulation of the most comprehensive catalog of global earthquake focal mechanisms available, spanning the years 1976 to the present. This chapter reviews this systematic analysis effort and some key characteristics of global seismicity as reflected in the CMT catalog. The review is not comprehensive, and I have limited the references to those directly associated with the evolution of the CMT Project. Section 4.18.2 provides a brief summary of the basic principles of the CMT method, as well as a description of how the algorithm has evolved since 1981. Section 4.18.3 presents a few representative maps and diagrams illustrating some of the primary characteristics of global seismicity. Section 4.18.4 describes recent discoveries of previously undetected earthquakes and suggests some possible future directions for the monitoring and study of global seismicity.

4.18.2 The CMT Method

The CMT method (Dziewonski et al., 1981) is based on the linear relationship that exists between the six independent elements of the zeroth-order moment tensor describing the earthquake source and the elastic vibrations of the Earth generated by the earthquake. This relationship holds as long as the earthquake source dimension is small relative to the wavelength of the seismic wave considered and the duration of the source is small relative to the period of the seismic wave. An additional component of the CMT algorithm is the consideration and estimation of the source centroid, which represents the spatial and temporal center of moment of the earthquake.

The CMT concept is general and can in principle be used with any recorded seismogram, given that the point-source approximation holds. A necessary condition is, of course, that synthetic waveforms corresponding to the observed seismograms can be generated with sufficient fidelity to allow a quantitative comparison. In the CMT algorithm, complete seismograms are compared in a least squares sense, requiring

that the phase and amplitude of the waveforms can be matched very well. In their original paper, Dziewonski et al. (1981) used normal-mode summation for a spherical Earth to generate the synthetic seismograms necessary to retrieve the moment-tensor elements and the earthquake centroid. They found that the body-wave portion of long-period ($T > 45$ s) seismograms could be fit well using spherical-Earth synthetic waveforms, but that surface waves, which are significantly distorted by propagation through Earth's laterally heterogeneous shallow structure, could not be fit well.

In a follow-up paper, Dziewonski and Woodhouse (1983) extended the CMT method to consider very long-period ($T > 135$ s) surface waves in the inversion for the moment tensor. Following the development of tomographic models of the Earth's upper mantle, the calculation of synthetic seismograms in the CMT processing was modified to include corrections for laterally varying Earth structure (Dziewonski et al., 1984; Woodhouse and Dziewonski, 1984). These corrections were particularly useful for improving the phase alignment of observed and synthetic waveforms and resulted in improved moment-tensor determinations.

Even with better tomographic models of the upper mantle, the prediction of intermediate-period (35–150 s) surface waves was not sufficiently good to allow direct comparison of observed and synthetic surface-wave phases at regional or teleseismic distances. However, following the development of global phase-velocity maps for Love and Rayleigh waves in this period range (e.g., Ekström et al., 1997), Arvidsson and Ekström (1998) extended the CMT algorithm to include intermediate-period surface waves in the inversion for source parameters. This development, in turn, allowed the CMT algorithm to be applied to earthquakes of a smaller magnitude. Since 2004, the routine CMT analysis has included the fitting of intermediate-period surface-wave seismograms in the analysis, significantly increasing the number of earthquakes that can be successfully analyzed each year (Ekström et al., 2012).

The routine determination of moment tensors has benefited from the increasing number of broadband seismographic stations deployed globally. The earliest year for which the CMT method has been applied systematically is 1976 (Ekström and Nettles, 1997). In that year, the experimental high-gain long-period (HGLP) seismographic network was operating, as were a small number of Standard Research Observatory (SRO) and International Deployment of Accelerometer (IDA) stations. In total, fewer than a dozen stations were available for any single event. With the development first of the Global Digital Seismographic Network (GDSN) operated by the US Geological Survey (USGS) and subsequently the GSN of the Incorporated Research Institutions for Seismology (IRIS) and the USGS, both the number of stations and the quality of the data have improved. These improvements have, in turn, led to better and more robust moment-tensor determinations. In our current CMT analysis, seismograms from more than 100 stations are routinely analyzed for each earthquake (Ekström et al., 2012).

4.18.3 Aspects of Global Seismicity

Figure 1 shows a map of focal mechanisms for shallow earthquakes in the CMT catalog for the period 1976–2012. While many focal mechanisms are obscured by focal mechanisms of later, nearby earthquakes, the map conveys key aspects of the pattern of global seismicity. The plate boundaries are clearly delineated by earthquake activity, except in areas of very slow plate motion, or near fast-spreading ridges, where the lithosphere is too weak to support earthquakes of significant magnitude. The consistency of focal mechanisms along the plate boundaries is also evident, particularly in subduction zones and along oceanic transforms. While some portions of the continents are devoid of $M > 5$ earthquakes, it is evident that over the 37-year period shown, many continental intraplate regions have experienced significant seismicity. The regional

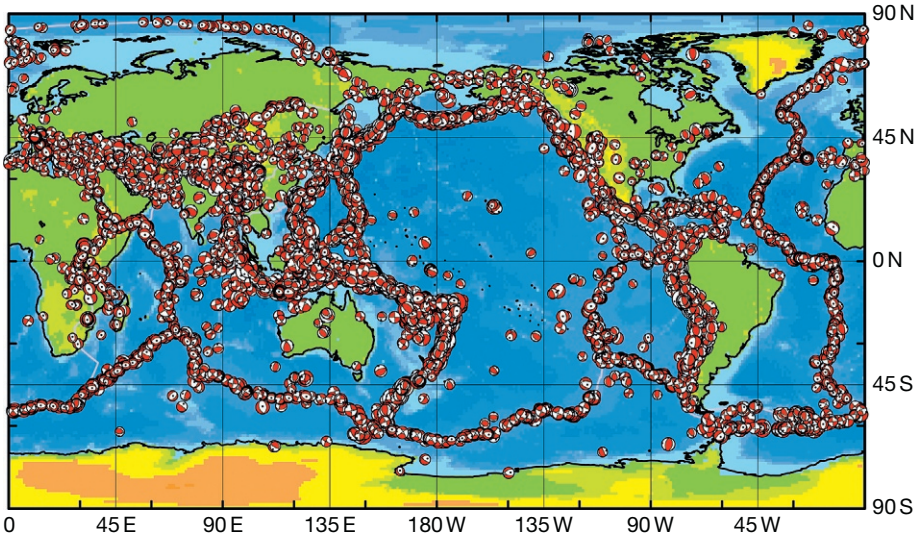


Figure 1 Map showing focal mechanisms for shallow earthquakes ($h < 70$ km) in the CMT catalog, 1976–2012. The earthquakes are shown in lower-hemisphere projection, and the diameter of the symbols increases linearly with the moment magnitude. Plate boundaries are shown but are nearly entirely covered by focal mechanisms.

variability of seismicity patterns (in terms of both frequency and the style of faulting) within oceanic plates is also striking.

Seismically active plate boundaries involving at least one oceanic plate are normally characterized by a relatively narrow zone of tectonic deformation, reflected in a narrow zone of earthquakes. Figure 2 shows focal mechanisms from the Alaska–Aleutian Island subduction zone, an area of very active seismicity and the site of several $M > 7.5$ earthquakes since 1976. The vast majority of the earthquakes are consistent with the shallow underthrusting of the Pacific Plate beneath the North American Plate. In the western portion of the subduction zone, the greater predominance of strike-slip focal mechanisms reflects the increasing obliquity of plate motion toward the west. In the eastern portion of the subduction zone, the pattern of seismicity is more distributed, with earthquakes associated with the plate collision occurring both in the interior of Alaska and within the Pacific Plate.

The most dramatic example of geographically distributed seismicity occurring in response to plate convergence is that associated with the collision of India and Eurasia. Figure 3 shows focal mechanisms for shallow earthquakes in this zone. The geographic association of earthquakes with topography is evident, and the zone of activity extends more than 1000 km north from the Himalayan front. Within the Tibetan Plateau, earthquake focal mechanisms primarily represent not the relative plate motion, but rather the strains associated with the gravity-controlled dynamics of the elevated plateau. The vast majority of focal mechanisms indicate strike-slip and normal faulting consistent with east–west extension.

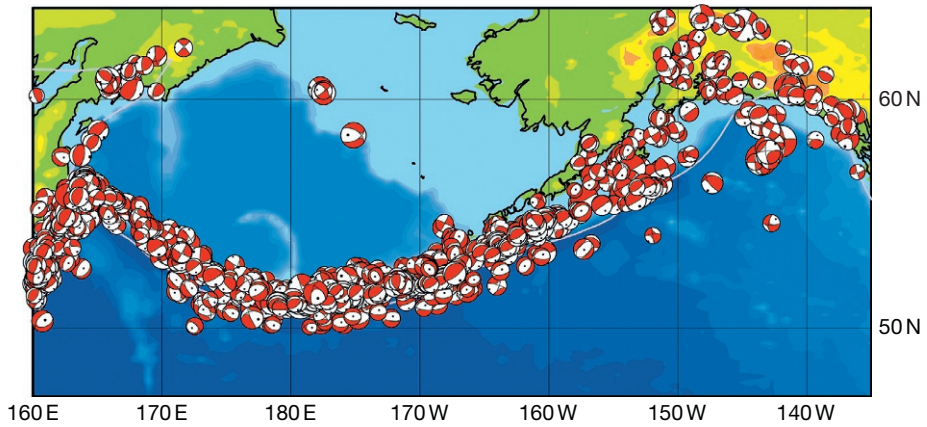


Figure 2 Map showing the Alaska–Aleutian subduction zone, with focal mechanisms for shallow earthquakes from the CMT catalog for the period 1976–2012. While seismic activity along the central and western part of the subduction zone is focused on the plate boundary, several intraplate earthquakes in the eastern zone indicate deformation within both the North American and the Pacific Plates.

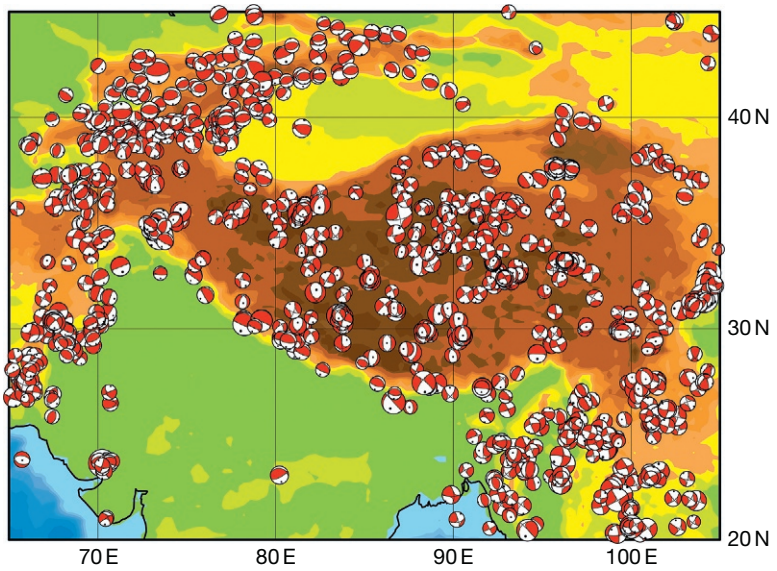


Figure 3 Map showing the India–Eurasia collision zone, with focal mechanisms for shallow earthquakes from the CMT catalog for the period 1976–2012. While earthquakes at the northern and southern edges of the Tibetan Plateau show mainly reverse motion associated with north–south compression, most of the earthquakes on the plateau reflect east–west extension.

Figure 4 shows the cumulative global moment of earthquakes since 1976. As is well known, most of the seismic moment (and corresponding strain) in a region occurs in the largest earthquakes, while small earthquakes, individually and collectively, account for only a small fraction of the cumulative moment. This division of cumulative seismic moment is illustrated by the contribution provided by earthquakes with $M < 6.5$ to the total, represented by the area shaded green in Figure 4. Over the 37-year period of the CMT catalog, 29 earthquakes have moment magnitudes of 8.0 or above. The largest earthquake to occur during the last 40 years is the 2004 Sumatra earthquake, which is easily identified in Figure 4 as the largest step in cumulative moment. In the figure, I have used the extended-source magnitude of 9.3, which was obtained in a CMT analysis that considered the spatial and temporal extent of the source (Tsai et al., 2005). The second-largest earthquake is the 2011 Japan earthquake. With a moment magnitude of 9.1 (Nettles et al., 2011), this event is also clearly identifiable as a step in cumulative moment in Figure 4.

Figure 5 shows the cumulative number of earthquakes in the CMT catalog since 1976. The main trend in this figure is the

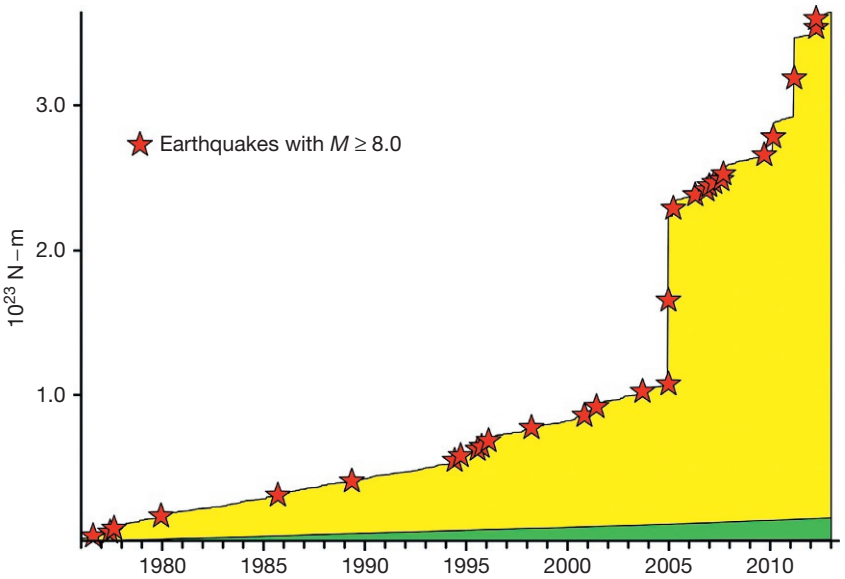


Figure 4 Cumulative moment of all earthquakes in the CMT catalog. Red stars indicate the times of earthquakes with moment magnitude $M \geq 8.0$. The field shaded yellow reflects the cumulative moment of earthquakes with $M \geq 6.5$. The green field reflects cumulative moment of earthquakes with $5.3 \leq M \leq 6.5$, and orange (barely visible) reflects earthquakes with $M \leq 5.3$. The contributions of the December 2004 Sumatra ($M = 9.3$) and March 2011 Japan ($M = 9.1$) earthquakes to the total cumulative moment are seen as the largest steps in the curve.

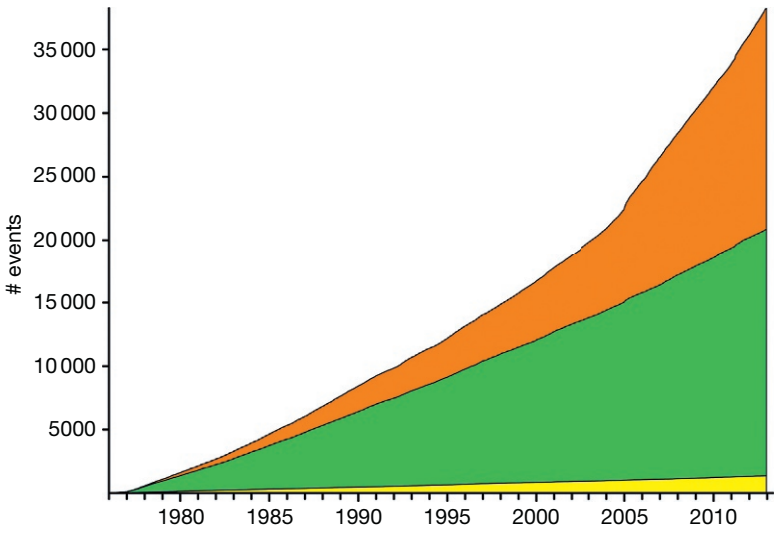


Figure 5 Cumulative number of earthquakes in the CMT catalog since 1976. Groups of earthquakes of different magnitudes are color-coded as in Figure 4 (orange, $M \leq 5.3$; green, $5.3 \leq M \leq 6.5$; and yellow, $M \geq 6.5$). An increase in the number of earthquakes in the catalog is seen for events with $M \leq 5.3$, especially since 2004 when intermediate-period surface waves were first included in the analysis.

gradual increase in the number of earthquakes with $M < 5.3$ that we have been able to analyze as a result of improvements in global station coverage and the CMT algorithm. In particular, the consideration of intermediate-period surface waves, which was initiated with earthquakes in 2004, nearly doubled the number of earthquakes that can be analyzed. Noticeable small increases in the seismicity rate can be observed associated with the many aftershocks that occurred in the months following the 2004 Sumatra and 2011 Japan earthquakes.

Most earthquakes occur along plate boundaries at shallow depths, and this is also where a maximum in cumulative seismic moment is observed. Figure 6 shows the depth distribution of cumulative seismic moment for the period 1976–2012. A distinct minimum in moment occurs around 350 km depth, and a local maximum is evident at the bottom of the upper mantle, around 600 km depth. While all deep earthquakes can be associated with current subduction or the likely location of recently subducted lithosphere, though not always unambiguously, the nature of seismic faulting at depth and its relationship to the physical state, stress, and strain of the mantle rock is not resolved.

Figure 7 shows an interesting example of seismic activity in the subducted slab beneath Chile and Argentina. The top panel of Figure 7 shows, in map view, the distribution of earthquakes along the plate interface and in the subducted Nazca slab. One of the deepest earthquakes is located to the east of the adjacent earthquakes, and in cross section (Figure 7, bottom), it is clear that a spatial separation exists between the main group of deep events and this 'outlier earthquake.' The outlier also exhibits a compression axis with a different orientation than the adjacent events, which have focal mechanisms consistent with downdip compression. The significance of outliers such as this one is hard to assess. If deep earthquakes are markers of subducted lithosphere, we must invoke some contortion of the subducting slab, or the existence of a second slab or slab fragment, in order to explain the existence of outlying earthquakes.

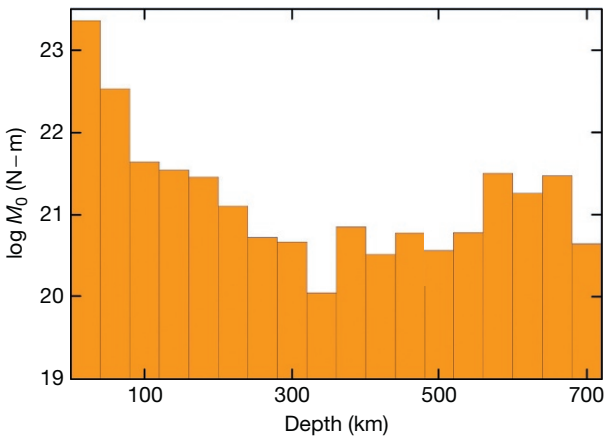


Figure 6 Graph showing the cumulative moment of earthquakes in the CMT catalog, 1976–2012, for different depths. The earthquakes have been binned in 40 km thick layers. Most of the cumulative moment is located in the top 40 km. A minimum occurs in the depth range 320–360 km. A local maximum occurs at the bottom of the upper mantle, in the range 560–680 km.

A further illustration of the complexity of deep seismicity is provided by the earthquakes beneath Fiji and Tonga, the world's most productive zone of deep seismicity. The top panel of Figure 8 shows the earthquakes in this region in map view; the bottom panel shows the seismicity projected onto a vertical plane with a strike close to that of the deep seismicity. In this view, it is clear that the earthquakes show complex spatial patterns, including both lineations in the seismicity and large volumes without significant seismicity. The deepest earthquakes, toward the south, are surrounded by a large area lacking significant earthquakes during the last 30 years. Although most of the compression axes shown in the bottom panel of Figure 8 are consistent with downdip compression, there are several clear and systematic deviations from this basic pattern, which require more detailed explanations in terms of the stresses and deformation in the subducted lithosphere.

4.18.4 Recent Discoveries and Future Directions

The study of global seismicity depends on the initial detection and location of seismic events. Several agencies are responsible for the monitoring of seismic activity on a global scale, in particular the National Earthquake Information Service (NEIS) of the USGS, the International Seismological Centre (ISC), and the International Monitoring System (IMS) in Vienna. All of these organizations rely primarily on high-frequency seismic signals for the detection of seismic events, leaving open the possibility that earthquakes deficient in radiated short-period energy may escape detection.

In recent work, we have applied a new method of array processing of continuous long-period vertical data from the GSN (Ekström, 2006; Ekström et al., 2003) to search for sources of long-period Rayleigh waves. The algorithm was applied to data filtered in the period range 40–150 s, in which seismic noise is low and knowledge of Earth's structure is sufficiently good to allow accurate deconvolution of surface-wave dispersion effects for arbitrary ray paths. In the initial application of the detection and location algorithm to data for the period 1993–2003, we found nearly 25 000 seismic sources with magnitudes $M > 4.5$, where the magnitude M is a good estimate of the moment magnitude. The vast majority of these events are standard earthquakes that are also found in the event catalogs of the NEIC, ISC, and the IMS. Approximately 5% of the events are, however, not listed in any of these catalogs and are therefore considered 'new.'

Figure 9 shows the locations of new earthquakes for the period 1993–2012. Most of the events occur in the southern hemisphere along the ridge-transform plate boundaries. Very few events are located in subduction zones. The events near plate boundaries may be earthquakes that have unusual rupture characteristics, or they may have gone undetected for other reasons (e.g., high microseismic noise or nodal radiation to many stations). Several other events are located in areas where earthquakes of this magnitude ($M > 4.5$) are unexpected or where they should have been detected if they represent standard earthquake sources. Most prominent in this category are the events on Greenland, which we have identified with a new phenomenon called 'glacial earthquakes' (Ekström et al., 2003, 2006), and

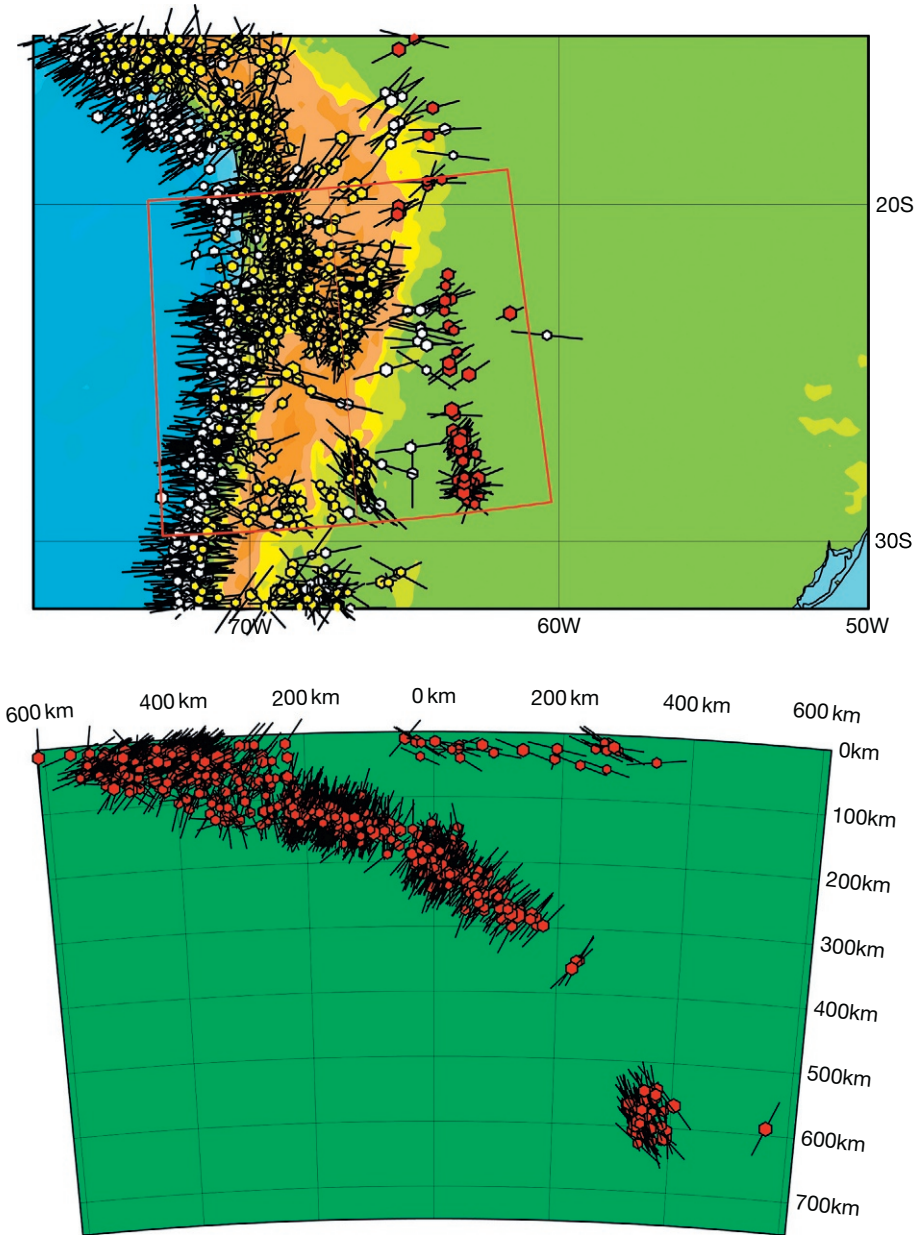


Figure 7 (Top) Compression axes of CMT moment tensors in the western part of South America, projected into the horizontal plane. The earthquakes are color-coded with respect to depth, with white symbols for shallow earthquakes and red symbols indicating the deepest events. The red box indicates the area shown in cross section in the bottom panel. The viewing direction is shown by the thin red line in the center of the box. (Bottom) Vertical cross section aligned with a viewing direction in the plane of the earthquakes associated with the subduction of the Nazca Plate beneath South America. Compression axes of CMT focal mechanisms are shown, projected into the plane of the cross section. At the bottom of the subduction zone, one outlying event has a compression axis that is rotated by 70° with respect to the downdip slab direction suggested by the main seismicity at this depth. This event is also displaced horizontally by nearly 200 km from the inferred location of the subducted slab.

we now understand to be the seismic expression of massive iceberg calving (Nettles et al., 2008). These events radiate surface waves with polarities and amplitudes consistent with a mass-sliding source, rather than with tectonic stress release. Figure 10 shows an example of a teleseismic record section from one of the glacial earthquakes, clearly demonstrating that signals for these events are large and can easily be identified at long periods. Several events in the Himalayas and in Alaska have been identified as catastrophic landslides (Ekström and Stark, 2013), and the events off the east coast of the United States may represent submarine landslides or slumps. Other groups of new events are spatially associated with volcanoes, and yet others lack, at this time, an obvious association or explanation.

These findings, and other recent discoveries of new events and processes that generate seismic radiation, indicate that the

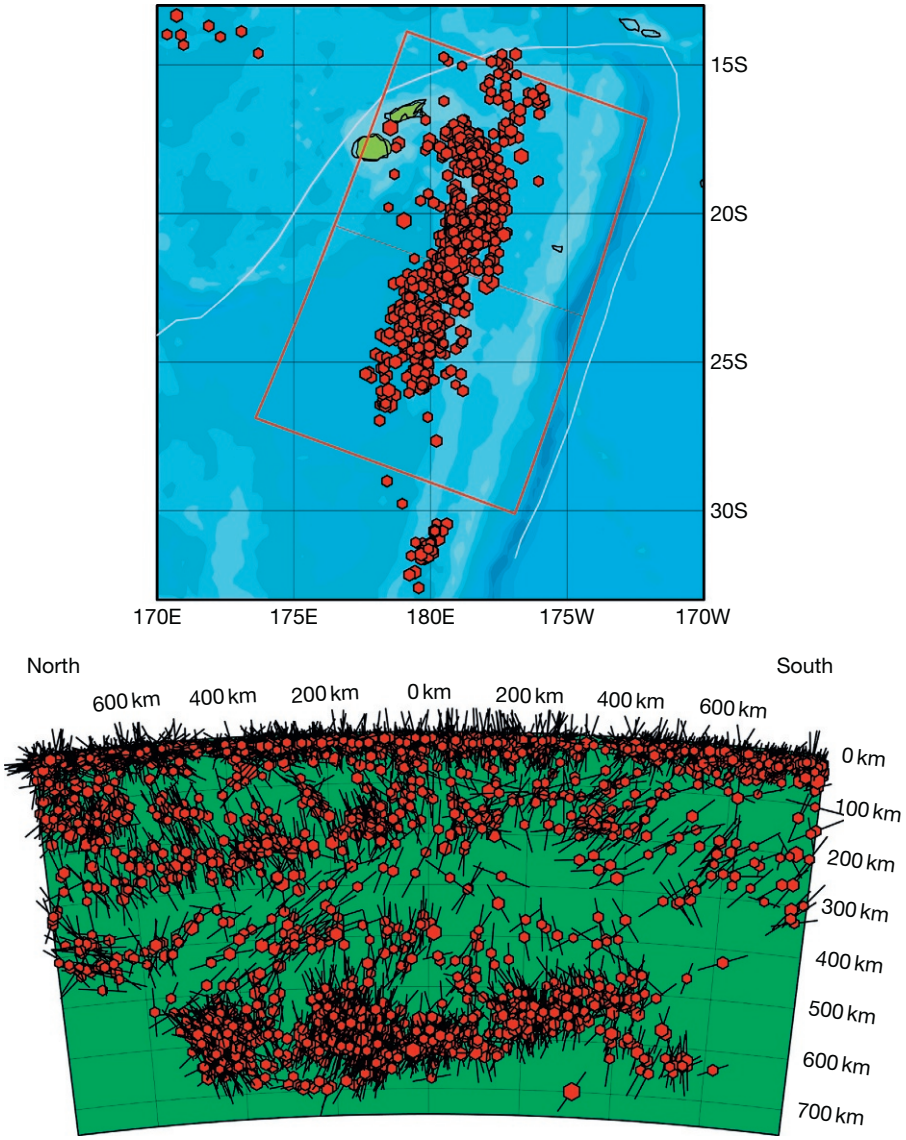


Figure 8 (Top) Map view of deep seismicity beneath Fiji–Tonga. The red box indicates the area viewed in cross section in the bottom panel. (Bottom) Vertical cross section aligned with a viewing direction perpendicular to the strike of the deep seismicity associated with the Tonga subduction zone. The compression axes of the focal mechanism are projected onto the vertical plane and show a very complex pattern of internal deformation in the plane of the slab. Some spatial trends in the locations and stress axes can be followed for hundreds of kilometers along the slab.

spectrum of seismic sources is richer than we had previously realized. These new events may be important for our understanding not only of lithospheric tectonics but potentially also of other phenomena such as magma motion and subaerial and submarine mass wasting. With the aid of the powerful infrastructure provided by the GSN, and by many new regional broadband seismographic networks, it seems likely that a focused effort to investigate the global seismic wave field in the period range 20–40 s for seismic events would be fruitful. Events with $M > 4.0$ routinely generate large surface waves in this period range, and it seems possible that, with recent advances in the characterization of crustal structure and shallow upper-mantle elastic and anelastic structure, these waves could be properly migrated and stacked for array processing and event detection.

Acknowledgments

I thank Adam Dziewonski, the founder and for many years the leader of the CMT Project, and Meredith Nettles, the current coleader with the author of the Global CMT Project. The seismic waveforms used in the CMT Project come primarily from stations of the Global Seismographic Network operated by IRIS and the USGS; the waveforms are distributed by the IRIS Data Management Center. Additional waveforms come from several other networks, including the Geoscope, Mednet, GEOFON, and the US and Canadian national seismic networks. We are grateful to everyone involved in the collection and distribution of data from these networks. The CMT catalog and associated results are accessible at the website www.globalcmt.org. The CMT Project has been continuously funded by the National

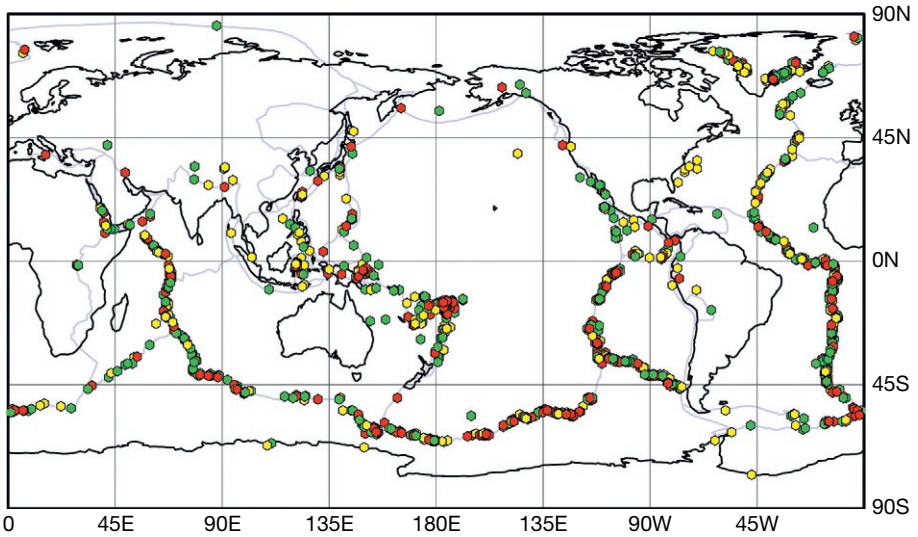


Figure 9 Map showing the locations of 3157 previously undetected earthquakes for the period 1993–2012 located using the surface-wave algorithm of Ekström (2006). The smallest earthquake has an estimated magnitude of $M = 4.6$ and the largest $M = 5.9$. The quality of detection is indicated by the color of the symbol, with red indicating the best quality, green very good quality, and yellow good quality.

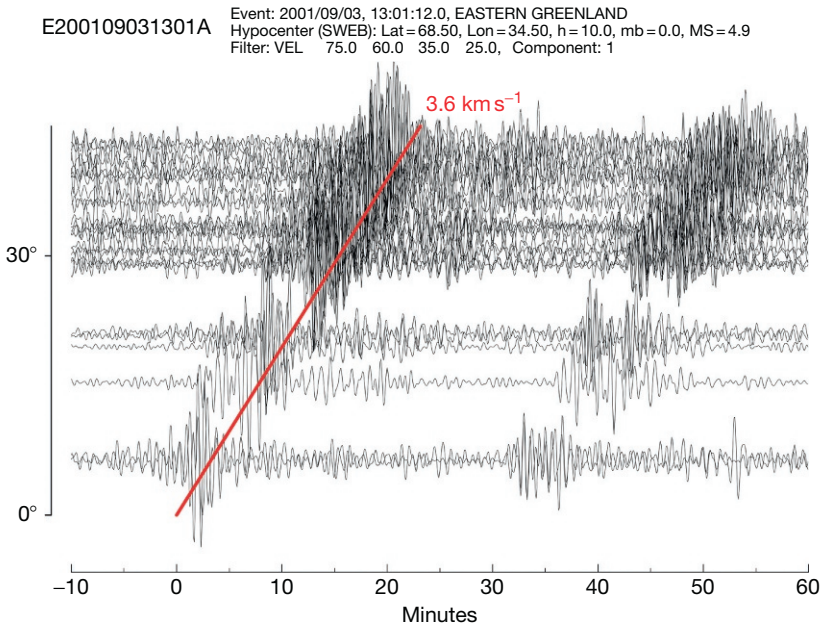


Figure 10 Record section for a glacial earthquake, showing the long-period Rayleigh wave arrivals at global stations out to a distance of 45°. Approximately 30 min after the main Rayleigh wave arrivals (which triggered the detection), a second arrival is seen with the same moveout, suggesting that a second, smaller event occurred in the same location.

Science Foundation since its inception, most recently by award EAR-1249167.

References

Arvidsson R and Ekström G (1998) Global CMT analysis of moderate earthquakes, $M_W \geq 4.5$, using intermediate period surface waves. *The Bulletin of the Seismological Society of America* 88: 1003–1013.

Dziewonski AM, Chou T-A, and Woodhouse JH (1981) Determination of earthquake source parameters from waveform data for studies of global and regional seismicity. *Journal of Geophysical Research* 86: 2825–2853.

Dziewonski AM, Franzen JE, and Woodhouse JH (1984) Centroid-moment tensor solutions for July-September 1983. *Physics of the Earth and Planetary Interiors* 34: 1–8.

Dziewonski AM and Woodhouse JH (1983) An experiment in the systematic study of global seismicity: Centroid-moment tensor solutions for 201 moderate and large earthquakes of 1981. *Journal of Geophysical Research* 88: 3247–3271.

Ekström G (2006) Global detection and location of seismic sources by using surface waves. *The Bulletin of the Seismological Society of America* 96: 1201–1212.

Ekström G and Nettles M (1997) Calibration of the HGLP seismograph network and centroid-moment tensor analysis of significant earthquakes of 1976. *Physics of the Earth and Planetary Interiors* 101: 219–243.

Ekström G, Nettles M, and Abers GA (2003) Glacial earthquakes. *Science* 302: 622–624.

Ekström G, Nettles M, and Dziewonski AM (2012) The Global CMT Project 2004–2010: Centroid-moment tensors for 13,017 earthquakes. *Physics of the Earth and Planetary Interiors* 200–201: 1–9.

Ekström G, Nettles M, and Tsai VC (2006) Seasonality and increasing frequency of Greenland glacial earthquakes. *Science* 311: 1756–1758.

Ekström G and Stark CP (2013) Simple scaling of catastrophic landslide dynamics. *Science* 339: 1416–1419.

Ekström G, Tromp J, and Larson EWF (1997) Measurements and global models of surface wave propagation. *Journal of Geophysical Research* 102: 8137–8157.

Gutenberg B and Richter CF (1941) Seismicity of the earth. *Geological Society of America Special Papers* 34: 131.

Gutenberg B and Richter CF (1949) *Seismicity of the Earth*. Princeton: Princeton University Press.

Nettles M, Ekström G, and Koss HC (2011) Centroid-moment tensor analysis of the 2011 off the Pacific coast of Tohoku earthquake and its larger foreshocks and aftershocks. *Earth, Planets and Space* 63: 519–523.

Nettles M, Larsen TB, Elósegui P, et al. (2008) Step-wise changes in glacier speed coincide with calving and glacial earthquakes at Helheim Glacier, Greenland. *Geophysical Research Letters* 35: L24503. http://dx.doi.org/10.1029/2008GL036127.

Tsai VC, Nettles M, Ekström G, and Dziewonski AM (2005) Multiple CMT source analysis of the 2004 Sumatra earthquake. *Geophysical Research Letters* 32: L17304. http://dx.doi.org/10.1029/2005GL023813.

Woodhouse JH and Dziewonski AM (1984) Mapping the upper mantle: Three dimensional modelling of earth structure by inversion of seismic waveforms. *Journal of Geophysical Research* 89: 5953–5986.

This page intentionally left blank

4.19 Tsunamis

K Satake, University of Tokyo, Tokyo, Japan

© 2015 Elsevier B.V. All rights reserved.

4.19.1	**Introduction**	478
4.19.2	**Recent Devastating Tsunamis**	478
4.19.2.1	The 2004 Indian Ocean Tsunami	478
4.19.2.2	The 2011 Tohoku Tsunami of Japan	480
4.19.3	**Other Examples**	482
4.19.3.1	Historical Tsunamis	482
4.19.3.1.1	The 1998 Papua New Guinea tsunami	482
4.19.3.1.2	The 1993 Hokkaido tsunami	482
4.19.3.1.3	The 1960 and 2010 Chilean tsunamis	483
4.19.3.1.4	The 1958 Lituya Bay tsunami	483
4.19.3.1.5	The 1946 Aleutian tsunami	483
4.19.3.1.6	The 1896 and 1933 Sanriku tsunamis	483
4.19.3.1.7	The 1883 Krakatoa tsunami	483
4.19.3.2	Prehistoric Tsunamis	483
4.19.3.2.1	The 1700 Cascadia tsunami	483
4.19.3.2.2	The seventeenth-century Hokkaido tsunami	484
4.19.3.2.3	The 869 Jogan tsunami	484
4.19.3.2.4	Tsunami from the Storegga Slide	484
4.19.3.2.5	Tsunami near the Cretaceous–Tertiary boundary	484
4.19.4	**Tsunami Observations**	484
4.19.4.1	Instrumental Measurements	484
4.19.4.1.1	Coastal tsunami measurements	484
4.19.4.1.2	Offshore tsunami measurements	485
4.19.4.1.3	Deep-ocean measurements	485
4.19.4.1.4	Satellite observations	485
4.19.4.1.5	Definition of amplitude and wave heights	486
4.19.4.2	Field Survey to Measure Tsunami Heights	486
4.19.4.2.1	Posttsunami surveys	486
4.19.4.2.2	Definitions of inundation and run-up heights	486
4.19.4.3	Study of Historical Documents	487
4.19.4.4	Geologic Methods	487
4.19.4.5	Tsunami Magnitude Scales	487
4.19.4.5.1	Imamura–Iida scale (m)	487
4.19.4.5.2	Tsunami intensity scale (I)	487
4.19.4.5.3	Tsunami magnitude (M_t)	487
4.19.4.6	Tsunami Catalogs	488
4.19.5	**Tsunami Generation**	488
4.19.5.1	Earthquake Tsunamis	488
4.19.5.1.1	Seafloor and water surface displacements	488
4.19.5.1.2	Fault parameters and seafloor displacement	488
4.19.5.1.3	Estimation of earthquake source parameters from tsunamis	488
4.19.5.1.4	Tsunamigenic and tsunami earthquakes	489
4.19.5.2	Tsunamis Caused by Landslides	490
4.19.5.2.1	Examples	490
4.19.5.2.2	Characteristics and discrimination of landslide tsunamis	490
4.19.5.2.3	Modeling landslide tsunamis	490
4.19.5.3	Volcanic Tsunamis	491
4.19.6	**Tsunami Propagation**	492
4.19.6.1	Weak Coupling between the Elastic Earth and Ocean	492
4.19.6.2	Ray-Theoretical Approach	492
4.19.6.3	Coastal Amplification and the Green's Law	492
4.19.6.4	Tsunamis as Gravity Waves	493
4.19.6.5	Basic Equations for Tsunami Propagation	493
4.19.6.6	Dispersion	494

 http://dx.doi.org/10.1016/B978-0-444-53802-4.00086-5

4.19.6.7	Numerical Computations	494
4.19.7	**Tsunami Warning Systems**	495
4.19.7.1	Far-Field Tsunami Warning Systems	495
4.19.7.2	Local Tsunami Warning Systems	496
4.19.8	**Tsunami Hazard Assessments**	497
4.19.8.1	Tsunami Hazard Maps	497
4.19.8.2	Probabilistic Tsunami Hazard Assessment	497
4.19.9	**Summary**	498
References		499

4.19.1 Introduction

A tsunami is a series of oceanic gravity waves generated by submarine or coastal geologic processes such as earthquakes, landslides, or volcanic eruptions. Most tsunamis are caused by the seafloor displacement due to large shallow earthquakes along subduction zones. In a rare case, an asteroid impact can also cause a tsunami. Tsunami is a Japanese word meaning 'harbor wave' but has been used internationally since the 2004 Indian Ocean tsunami. Tsunamis are usually small and hardly noticeable when forming in the deep ocean but become large and can cause major damage when they approach coasts or harbors. The recent 2011 Tohoku tsunami of Japan showed the type of destruction power a tsunami can have (**Figure 1**). The 2004 Indian Ocean tsunami demonstrated that a tsunami can travel at transoceanic distances and cause damage far from its source.

Tsunamis have many different aspects that are studied by researchers in various fields (Bernard and Robinson, 2009). Their generation is mostly related to geologic processes including earthquakes (**Chapter 4.01**). The propagation and offshore observations of a tsunami are the subjects of oceanography, and basic hydrodynamics are needed to understand the characteristics of a tsunami. Coastal behaviors, such as run-up on beaches or resonance in bays, are, for the most part, studied by coastal engineers. Officials in local and national governments and civil defense and emergency managements are responsible for conducting operational tsunami warnings and mitigating tsunami disasters.

Figure 1 Tsunami inundation in the Sendai plain following the 2011 Tohoku earthquake. Reproduced with permission from Mainichi Newspaper.

4.19.2 Recent Devastating Tsunamis

4.19.2.1 The 2004 Indian Ocean Tsunami

A giant earthquake ($M_w = 9.1$, according to the US Geological Survey (USGS)) occurred off the coast of Sumatra Island on 26 December 2004 (Lay et al., 2005; Stein and Okal, 2005; **Chapter 4.18**). This earthquake generated a tsunami that devastated the shores of the Indian Ocean (**Figure 2**); more than 200 000 people lost their lives. The total number of victims, dead and missing, was the largest in Indonesia (163 795), followed by the disasters in Sri Lanka (35 399), India (16 389), and Thailand (8345) (International Federation of Red Cross and Red Crescent Societies, 2005). Within 30 min of the earthquake, the tsunami devastated Banda Aceh and other coastal villages of Sumatra Island (Lavigne et al., 2009). The tsunami reached the coasts of Thailand, Sri Lanka, and India in approximately 2 h after the earthquake. The tsunami propagated farther and reached the east coast of Africa several hours after the earthquake and caused 298 casualties in Somalia.

To document the 2004 tsunami, many scientists from all over the world visited the affected coasts (Borrero, 2005; Lavigne et al., 2009; Liu et al., 2005; Satake et al., 2006; Tsuji et al., 2006). The measured tsunami heights near Sumatra Island, particularly in the Aceh province, were for the most part larger than 20 m with a maximum height above 30 m (**Figure 3**). The tsunami heights along the Andaman Sea coast varied greatly: 5–15 m near Thailand but less than 3 m near Myanmar. The tsunami heights were up to 5 m near India's Andaman Islands. In Sri Lanka, the tsunami heights were 5–15 m. The height distribution of the 2004 Indian Ocean tsunami is consistent with its damage distribution and indicates that the source of the large and damaging tsunami was concentrated in the southern 700 km section of the aftershock zone. Sedimentologic characteristics of the modern tsunami deposits were also examined (Jaffe et al., 2006; Moore et al., 2006) with hopes that they provide valuable information for the studies of paleotsunami deposits (**Section 4.19.4.4**).

The tsunami was instrumentally recorded by coastal tide gauges in the Indian Ocean (Merrifield et al., 2005; Nagarajan et al., 2006; Rabinovich and Thomson, 2007) and in the Atlantic and Pacific Oceans (Rabinovich et al., 2011; Titov et al., 2005a). This tsunami was also recorded by other instruments, some of which were used for the first time. The propagation of the tsunami in deep water was captured by a satellite altimeter (Gower, 2005; Smith et al., 2005). The signal of the tsunami was also recorded on the horizontal components of broadband seismic stations near the coasts of the affected areas (Yuan et al., 2005) and on hydrophones (Hanson et al., 2007).

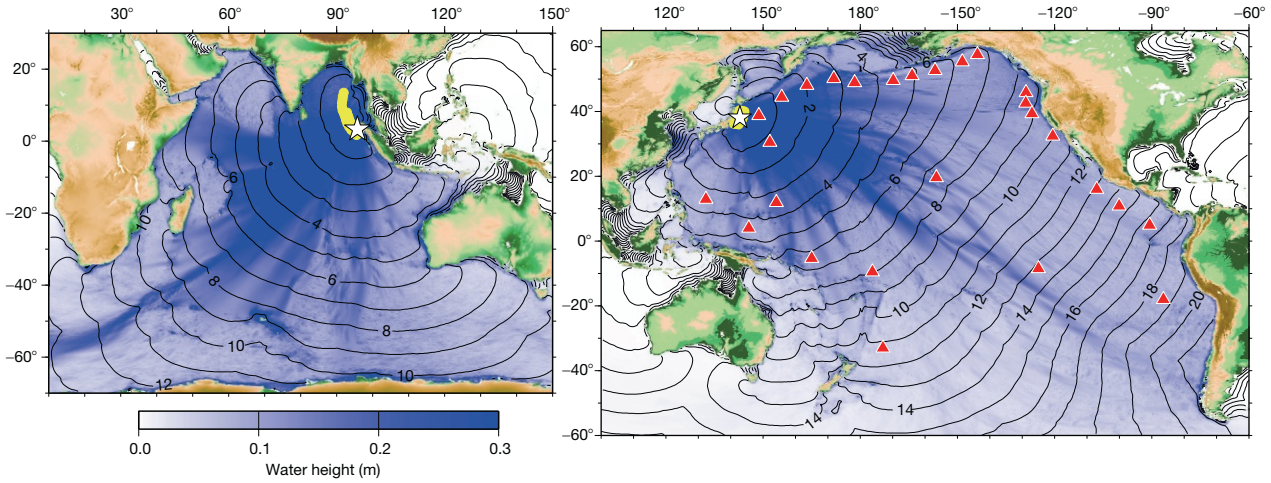


Figure 2 Tsunami propagation computed from the 2004 Sumatra–Andaman earthquake (left) and the 2011 Tohoku earthquake (right). The curves indicate wave fronts, which are plotted in 1-h increments from the tsunami source (yellow area). White stars indicate the epicenters of the mainshock (data from USGS). The blue color indicates the computed maximum water height (or energy) distribution. The red triangles in the right panel show the locations of the ocean-bottom pressure gauge stations (deep-ocean assessment and reporting of tsunamis; see **Section 4.19.4.1.3**).

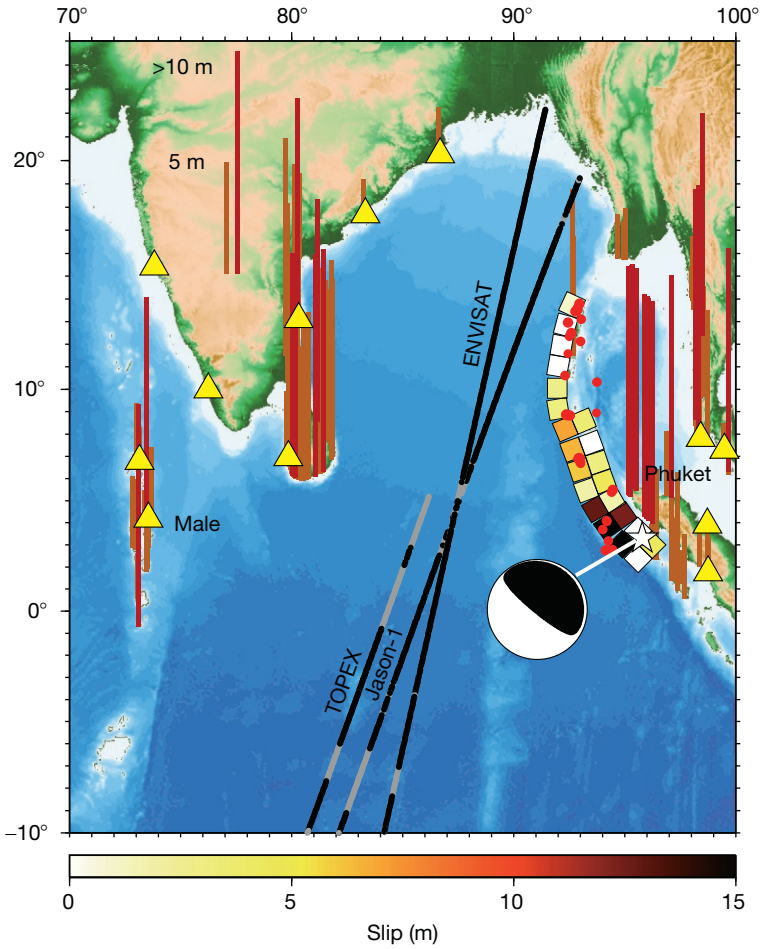


Figure 3 Locations of the instrumental observations for the 2004 Indian Ocean tsunami. Tide gauge stations are indicated by white triangles and satellite paths are indicated by solid lines. The colored squares, which are represented by the scale below the map, show the slip distribution on the fault for the 2004 Sumatra–Andaman earthquake via an inversion of tsunami waveforms recorded on tide gauges and sea-surface heights measured by satellites (Fujii and Satake, 2007). The tsunami heights measured by field surveys are shown as orange (<10 m) and red (>10 m) bars (NOAA NGDC Tsunami Database). The red circles and beach ball show the 1-day aftershocks (USGS) and focal mechanism (the Global CMT Project, see **Chapter 4.18**).

Seismological analyses of the Sumatra–Andaman earthquake indicate that the source region extended from off the coast of Sumatra through the Nicobar and Andaman Islands with a total length of 1200–1300 km, and the total duration of the rupture was about 500 s (Ishii et al., 2005; Tsai et al., 2005). The fault slip was largest off the coast of northern Sumatra, followed by the coasts of the Nicobar Islands (Ammon et al., 2005). The fault slip around the Andaman Islands was estimated to be small from seismological analysis and was speculated to be slow (Bilham, 2005; Lay et al., 2005).

Satellite image analyses indicate that sea-level changes occurred in Sumatra, the Nicobar Islands, and the Andaman Islands, suggesting that the source length was as long as 1600 km (Meltzner et al., 2006; Tobita et al., 2006). Field investigations indicate that the largest subsidence was 3 m and occurred at Great Nicobar, while the small islands off the west coast of Middle Andaman were uplifted about 1.5 m (Kayanne et al., 2007; Malik and Murty, 2005). The tide gauge record for Port Blair in South Andaman showed a slow subsidence with a duration between 15 and 40 min (Singh et al., 2006).

Using tsunami arrival times, the length of the tsunami source was estimated to be 600–700 km (Lay et al., 2005) but was later revised to be ~900 km (Neetu et al., 2005), which extended its length to the southern Andaman Islands. The satellite altimeter data support a longer length of more than 1000 km for the tsunami source (Fine et al., 2005b; Hirata et al., 2006; Song et al., 2005). The tide gauge data do not support a tsunami source that extended to the Andaman Islands (Fujii and Satake, 2007; Tanioka et al., 2006); however, Piatanesi and Lorito (2007) estimated that a slip occurred on the deeper portion of the slab, east of the Andaman Islands. Joint inversion of tide gauge and satellite data supports a source length of ~900 km (Fujii and Satake, 2007). Seno and Hirata (2007) speculated that the fast rupture of the source was followed by slow slip at the shallower portion near the trench.

Historical and paleoseismological studies indicate that great earthquakes have occurred in the past in the Indian Ocean. The source of the 1881 earthquake near Car Nicobar was studied using the tsunami waveforms recorded by tide gauges around the Bay of Bengal (Ortiz and Bilham, 2003). The potential hazard of a tsunami in the Indian Ocean was reported prior to the 2004 tsunami via a numerical simulation of the 1833 earthquake off the coast of Sumatra (Cummins, 2004). Paleoseismological studies of tsunami deposits have been conducted since 2004 in Sumatra Island (Monecke et al., 2008), Thailand (Fujino et al., 2009; Jankaew et al., 2008; Prendergast et al., 2012), the Andaman and Nicobar Islands (Malik et al., 2011; Rajendran et al., 2013), and India (Rajendran et al., 2006), and the results prove that tsunamis have occurred in the past. A penultimate earthquake was estimated to have occurred around AD 1300–1450 in Thailand (**Figure 4**), AD 1290–1400 in Sumatra, AD 1250–1450 near the Andaman and Nicobar Islands, post AD 1600 in South Andaman Island, and around AD 1020–1160 along the Indian coast. These various dates may indicate that the penultimate earthquake was not exactly the same as the 2004 Sumatra–Andaman earthquake.

Figure 4 Sand deposited north of Phuket, Thailand, by the 2004 Indian Ocean tsunami (A) and by a previous tsunami in the fourteenth or fifteenth century AD (B). Modified from Jankaew K, et al. (2008) Medieval forewarning of the 2004 Indian Ocean tsunami in Thailand. *Nature* 455: 1228–1231, with permission.

4.19.2.2 The 2011 Tohoku Tsunami of Japan

A giant earthquake occurred off the northern coast of Honshu, Japan, on 11 March 2011. This magnitude 9.0 event was the largest earthquake in the history of Japan and caused a devastating tsunami disaster (**Figure 1**) and serious damage to the nearby Fukushima Daiichi Nuclear Power Station. The earthquake and subsequent tsunami were the cause of approximately 15 900 deaths, 3100 missing bodies, and 6000 injured persons, with 92% of the casualties being the result of drowning due to the tsunami, according to the Central Disaster Management Council of Japan. The earthquake was officially named 'The 2011 off the Pacific coast of Tohoku Earthquake' by the Japan Meteorological Agency (JMA) but is often abbreviated as the '2011 Tohoku (-oki) earthquake' in scientific literature and in this section. The disaster, which included the Fukushima accident, was officially named 'the Great Eastern Japan Earthquake Disaster' by the Japanese Cabinet.

The 2011 Tohoku earthquake occurred at the boundary between the Pacific plate and the overlying plate. The Pacific plate subducts beneath northern Honshu along the Japan Trench at a rate of approximately 8 cm per year. This subduction causes westward movement and an east–west compression of northern Honshu, as observed prior to 2011 by the land-based Global Positioning System (GPS) network.

The focal mechanism solution for the 2011 Tohoku earthquake shows thrust-type fault movement on a shallowly dipping plane, indicating that it was a typical interplate earthquake. The nationwide GPS network of Japan recorded large movements in directions opposite to the previous observations during the interseismic period, with maximum motions of 5.3 m eastward and 1.2 m downward (Ozawa et al., 2011). The repeated marine geodetic measurements, conducted via onboard GPS and acoustic signals (GPS/A), showed that the seafloor near the epicenter moved 31 m horizontally and 3 m vertically (Kido et al., 2011; Sato et al., 2011). The offline ocean-bottom pressure gauges also recorded a very large horizontal displacement of 60 m and a vertical displacement

of 5 m (Ito et al., 2011). Repeated multibeam sonar soundings indicated that the seafloor moved approximately 50 m horizontally near the trench axis (Fujiwara et al., 2011).

The 2011 tsunami was first recorded by offshore gauges (**Figure 5**). With the ocean-bottom pressure gauge TM1 (located at about 76 km off the coast of Sanriku, at a water depth of 1600 m, and connected to a seafloor cable), a water rise of ~2 m was recorded immediately after the earthquake, followed by an impulsive wave that caused an additional 3 m rise within 2 min (Fujii et al., 2011). A nearshore GPS wave gauge, located about 10 km offshore at a water depth of 200 m, recorded similar tsunami waveforms with larger amplitudes 12 min after the earthquake. The arrival of the tsunami was recorded with a tide gauge at Kamaishi Port about 30 min after the earthquake, but the large tsunami destroyed the tide gauge system and the maximum amplitude was not recorded. The offshore tsunami generation and its propagation toward the Sanriku coast was thus recorded by the offshore ocean-bottom pressure gauge and GPS wave gauge, and the data were transmitted to the JMA in real time and were used to update the tsunami warnings (see **Section 4.19.7.2**).

The tsunami was also recorded by deep-ocean-bottom tsunami gauges of DART (operated by the National Oceanic and Atmospheric Administration (NOAA) and other agencies) and coastal gauges throughout the Pacific Ocean (**Figure 2**). The maximum trough to crest tsunami heights were 3–4 m on the west coast of the United States and along the southern coast of Chile (Heidarzadeh and Satake, 2013). The tsunami caused two fatalities outside of Japan: one in the United States and one in Indonesia.

The heights of the 2011 Tohoku tsunami near the Japanese coast were measured by the Joint Survey Group, which consisted of ~300 researchers from more than 60 organizations (Mori et al., 2011). The total number of measurement points exceeded 5900 (**Figure 5**). The tsunami reached the Sanriku coast about 30 min after the earthquake, with maximum heights of almost 40 m. The surveys show that the tsunami heights varied locally, particularly along the sawtooth-shaped Sanriku coast. The tsunami reached the Sendai plain about 1 h after the earthquake and inundated regions as far as 5 km from the coast. The total area of the 2011 tsunami inundation was estimated to be 561 km^2 by the Geospatial Information Authority of Japan.

The tsunami left sand and mud deposits in the Sendai plain and provided a valuable opportunity to study the sedimentologic characteristics of modern tsunami deposits (Goto et al., 2011). In the Sendai plain, a tsunami deposit from a previous earthquake, the AD 869 Jogan earthquake (see **Section 4.19.3.2.3**), was also found and recorded (Minoura et al., 2001; Sawai et al., 2012). While the 2011 tsunami inundated nearly 5 km from the coast, sand deposits with thicknesses of more than 0.5 cm can be traced up to 3 km, and only mud was deposited between 3 and 5 km (Goto et al., 2011). This indicates that previous studies on tsunami deposits and earthquake modeling, assuming that the extent of a tsunami deposit is the same as the inundation area, may

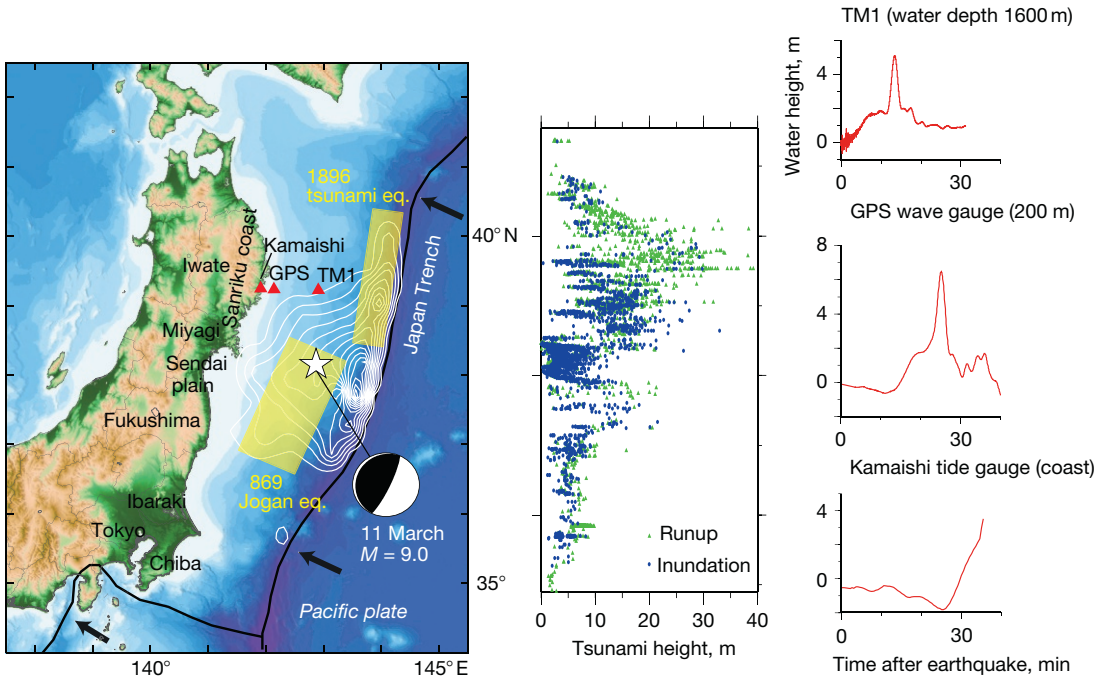


Figure 5 (Left) The source region of the 2011 Tohoku earthquake (the epicenter is shown by the white star). The red triangles show the locations of the bottom pressure, GPS wave, and coastal tide gauges. The plate boundaries and the relative motions are shown by thick black curves and arrows, respectively. The proposed fault models (yellow rectangles) of the 1896 Sanriku tsunami earthquake (Tanioka and Satake, 1996a) and the AD 869 Jogan earthquake (Sawai et al., 2012) are shown. The beach ball shows the focal mechanism solution of the mainshock (USGS). The slip distribution estimated from the tsunami waveforms is shown by contours with 4 m intervals (Satake et al., 2013a). (Center) The tsunami run-up and inundation heights compiled by the Joint Survey Group (Mori et al., 2011). (Right) The tsunami waveforms recorded by the bottom pressure gauge, GPS wave gauge, and coastal tide gauge (Fujii et al., 2011).

have underestimated the sizes of earthquakes (Namegaya and Satake, 2014; Sugawara et al., 2013).

Tsunami models indicate that the 2011 earthquake was similar to a combination of the 1896 Sanriku 'tsunami earthquake' and a 'Jogan-type' deeper interplate earthquake (Fujii et al., 2011; Satake et al., 2013a). A huge slip near the trench axis, similar to the 1896 Sanriku tsunami earthquake, caused the impulsive tsunami waves recorded by the bottom pressure and GPS wave gauges and was responsible for the high tsunamis along the Sanriku coast (**Figure 5**). The fault motion and large slip along the deeper plate interface, similar to a previous model of the AD 869 Jogan earthquake, produced a long-wavelength seafloor displacement and caused the first gradual rise in seawater at the location of the gauges and the large tsunami inundation in the Sendai plain.

4.19.3 Other Examples

In this section, other examples of tsunamis are briefly described in reverse chronological order. Their sources are shown in Figure 6. The most recent tsunamis that were instrumentally recorded are presented first, followed by prehistoric tsunamis.

4.19.3.1 Historical Tsunamis

4.19.3.1.1 The 1998 Papua New Guinea tsunami

A moderate earthquake (M_w=7.1) along the northern coast of New Guinea generated a tsunami on 17 July 1998, whose run-up heights were as large as 15 m around Sissano Lagoon, which is near the epicenter. More than 2000 casualties were reported. The large tsunami (>5 m), however, was limited to a small segment (~40 km) of the coast (Kawata et al., 1999), and the amplitudes recorded by the tide gauges around Japan were <10 cm (Satake and Tanioka, 2003), suggesting that the effects of the tsunami were local. Posttsunami marine surveys revealed bathymetry features that focused the energy of the tsunami toward the Sissano Lagoon and that suggest a possible source of submarine mass movements (Matsuyama et al., 1999; Tappin et al., 1999). Numerical simulations showed that a secondary source, in addition to the earthquake fault motion, which may have been caused by the submarine landslide, was needed to reproduce the large local tsunami heights (Geist, 2000; Satake and Tanioka, 2003; Synolakis et al., 2002; Titov and Gonzalez, 2001). The deposit of this tsunami was studied by Gelfenbaum and Jaffe (2003).

4.19.3.1.2 The 1993 Hokkaido tsunami

The largest tsunami disaster in Japan prior to the 2011 Tohoku event occurred in the Sea of Japan and was caused by the southwest Hokkaido earthquake (M_w=7.8) on 12 July 1993. The tsunami attacked Okushiri Island within a few minutes after strong ground shaking and caused more than 200 casualties. The tsunami heights ranged from 5 to 10 m around the island with a maximum tsunami run-up of more than 30 m near a small valley close to Monai (Shuto and Matsutomi, 1995). The deposit of this tsunami was studied (Nanayama et al., 2000; Nishimura and Miyaji, 1995). Numerical tsunami computations were conducted to investigate the source model and to reproduce the measured tsunami heights, tsunami waveforms recorded by the tide gauges, and seismic or geodetic

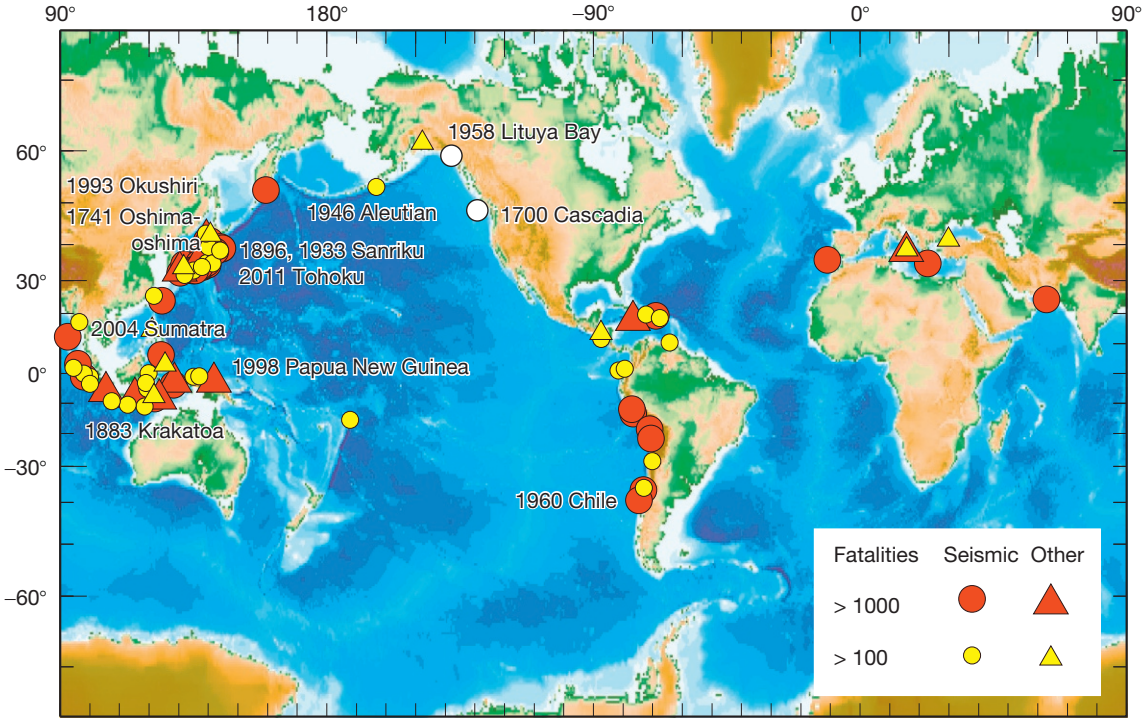


Figure 6 Distribution of tsunami occurrences throughout the world. Symbols represent the historical events with intensity ≥2 and validity ≥3 from the NOAA NGDC Tsunami Database.

data (Takahashi et al., 1995; Tanioka et al., 1995; Titov and Synolakis, 1997).

4.19.3.1.3 The 1960 and 2010 Chilean tsunamis

The largest earthquake ($M_w = 9.5$) of the twentieth century occurred off the shore of southern Chile on 22 May 1960. The earthquake and tsunami caused damage on the Chilean coast and resulted in more than 1000 casualties. While historical documents show that the previous large earthquakes occurred in 1837, 1737, and 1575, geologic evidence for coseismic subsidence and tsunami was found only for the 1575 and 1960 events (Cisternas et al., 2005). The 1960 tsunami propagated across the Pacific Ocean. About 15 h after the earthquake, the tsunami reached the Hawaiian Islands and caused 61 deaths. The tsunami continued across the Pacific, reaching Japan in about 23 h, and caused 139 casualties. The large (~5 m) tsunami that reached Japan was attributed to focusing and resonance effects. Because of the sphericity of the Earth and the bathymetric features of the Pacific Ocean, the energy of the tsunami was focused in the direction of the Japanese islands (Satake, 1988; Shuto, 1991). Some bays along the Pacific coast of Japan have a shape whose characteristic period coincided with the predominant period (about 1 h) of the transpacific tsunami (Abe, 2000). This tsunami motivated the formation an international tsunami warning system for the Pacific.

In 2010, another great ($M_w = 8.8$) earthquake, known as the Maule earthquake, occurred just to the north of the 1960 source area. The tsunami from this earthquake affected a segment of approximately 800 km along the Chilean coast with a local maximum run-up height of 29 m, caused structural damage and more than 100 casualties, and created tsunami deposits and other geologic features (Fritz et al., 2011; Mas et al., 2012; Morton et al., 2011). The effects of the tsunami along the coast were prolonged due to the trapping of energy on the continental shelf, which is known as edge waves (Fujii and Satake, 2013; Yamazaki and Cheung, 2011). The tsunami also propagated across the Pacific Ocean and was recorded on many DART, bottom pressure, GPS wave, and tide gauges (Fujii and Satake, 2013; Kato et al., 2011; Saito et al., 2010). The observed arrival of the transpacific tsunami was as late as 30 min compared to the theoretical arrival time (see **Section 4.19.6.1**).

4.19.3.1.4 The 1958 Lituya Bay tsunami

The largest water run-up was recorded in Lituya Bay, Alaska. A large strike-slip earthquake ($M_w = 7.9$) on 10 July 1958 triggered a huge rockslide in this 10 km long bay. The slide was 700–900 m wide, 1000 m long, and 90 m thick, resulting in a total volume of about 3×10^7 m^3. The rockslide generated water waves that surged up the opposite slope and stripped the trees up to 520 m altitude (Fritz et al., 2009; Lander and Lockridge, 1989; Miller, 1960). The large waves were confined to the bay, which is why this case is not considered a typical tsunami.

4.19.3.1.5 The 1946 Aleutian tsunami

The Aleutian earthquake that occurred on 1 April 1946 was very unusual. The earthquake generated huge tsunamis near the Aleutians and Hawaii, even though its surface wave magnitude (M_S) was only 7.4. This is a typical example of a 'tsunami earthquake,' which generates much larger tsunamis than expected from its seismic waves (Kanamori, 1972; see **Section 4.19.5.1.4**). The Scotch Cap Light on a 10 m cliff on Unimak Island, about 100 km from the epicenter, was totally demolished by the tsunami (Lander and Lockridge, 1989). The tsunami traveled southward through the Pacific Ocean and reached the Hawaiian Islands by the morning of the April Fool's Day. Without any warning, the tsunami took 159 lives in the Hawaiian Islands. A Pacific tsunami warning system was introduced after this tsunami. Recent studies indicate that a slow earthquake rupture process and a large submarine landslide were possibly responsible for the far-field and near-field tsunamis, respectively (Okal et al., 2002, 2003).

4.19.3.1.6 The 1896 and 1933 Sanriku tsunamis

The 15 June 1896 Sanriku earthquake is another example of a 'tsunami earthquake' (Kanamori, 1972; Tanioka and Satake, 1996a). The earthquake ($M_S = 7.2$) was not great in size and only caused weak ground shaking (2–3 on the Japanese seismic intensity scale, corresponding to IV–V on the Modified Mercalli Intensity scale). However, a tsunami known as the 'Meiji Sanriku tsunami' was very destructive; the run-up height was >10 m with a maximum value of 38 m and affected a 200 km segment along the Sanriku coast. The tsunami resulted in 22 000 casualties and was the worst tsunami disaster in Japan in the last few centuries.

The Sanriku coast also experienced a tsunami on 2 March 1933, known as the 'Showa Sanriku tsunami,' which had a maximum height of 24 m and caused about 3000 casualties. This tsunami originated from a large normal-fault-type earthquake ($M = 8.2$) in the outer-rise region (Kanamori, 1971).

4.19.3.1.7 The 1883 Krakatoa tsunami

The 1883 eruption of Krakatoa (Indonesia) is one of the largest volcanic eruptions recorded in human history. The sound waves generated by the eruption were heard from a distance of over 4000 km across the Indian Ocean. The eruption refaced the island of Krakatoa; the northern portion of the island was replaced by a caldera up to 270 m deep. The eruption also generated tsunamis that destroyed nearby coastal villages and took about 34 000 lives (Simkin and Fiske, 1983). The maximum tsunami height along the Sunda Strait was about 15 m and the tsunami waveform was recorded by a tide gauge at Batavia (now Jakarta). The tsunami generation mechanism from various types of volcanic activities, such as caldera collapses, submarine explosions, and pyroclastic flows, has been proposed (Maeno and Imamura, 2011; Nomanbhoy and Satake, 1995; Yokoyama, 1987). Sea-level disturbances were observed by tide gauges in India, in New Zealand, and even at San Francisco, but those observed at very far places are interpreted as sea waves coupled by air waves from the eruption (Pelinovsky et al., 2005; Press and Harkrider, 1966).

4.19.3.2 Prehistoric Tsunamis

4.19.3.2.1 The 1700 Cascadia tsunami

The Cascadia subduction zone along the Pacific Northwest of the United States and Canada has produced no great earthquakes since the region's written history began in the 1800s.

Geologic studies show the evidence of past earthquakes and tsunamis. Evidence of a great prehistoric earthquake that occurred along the Cascadia subduction zone has been found both near and far from the source (Atwater et al., 2005). Along the Pacific coast of North America, geologic and tree-ring evidence indicates that coseismic subsidence took place about 300 years ago. In addition, tsunami deposits found throughout the Pacific Northwest indicate that this earthquake generated a sizable tsunami. In Japan, historical records show that a widespread tsunami with heights of 1–5 m struck the Pacific coast. From these records, the date and size of the earthquake were inferred to be 26 January 1700 and $M_w = 8.7–9.2$, respectively (Satake et al., 1996, 2003).

4.19.3.2.2 The seventeenth-century Hokkaido tsunami

Historical documents are limited in the last 150 years for Hokkaido, which is the northernmost island of Japan. Tsunami deposits show that the southern Kuril trench repeatedly produced tsunamis larger than those recorded in the last few centuries. Prehistoric sand sheets extend as much as 3 km inland along the 200 km segment of Hokkaido's Pacific coast, where the recent tsunamis from 1952 to 2003 Tokachi-oki or 1894 and 1973 Nemuro-oki earthquakes ($M_w \sim 8$) penetrated less than 1 km from the coast. The time intervals between the extensive sand sheets average about 500 years, and the most recent event occurred in the seventeenth century. These anomalous tsunamis, characterized by a large inundation area and a long recurrence interval, are best explained by earthquakes that ruptured multiple segments, including the Tokachi-oki and Nemuro-oki regions of the Kuril subduction zone (Nanayama et al., 2003; Satake et al., 2008; Sawai et al., 2009).

4.19.3.2.3 The 869 Jogan tsunami

The predecessor of Japan's largest earthquake, the 2011 Tohoku earthquake, is considered to be the 869 Jogan earthquake. The Japanese national history book 'Nihon Sandai Jitsuroku' depicts strong shaking, ground fissures, collapsed houses, and numerous casualties due to the earthquake. It also describes the tsunami to have produced a sound like thunder and inundated along the river up to a castle. The flooded area is reported to be hundreds of kilometers, causing 1000 people to drown. In addition, tsunami deposits were found in geologic layers in coastal lowlands, more than 4 km from the current coast in the Sendai plain (Minoura and Nakaya, 1991; Namegaya and Satake, 2014; Sawai et al., 2008). From the distribution of the tsunami deposits and computed inundation area, the size ($M = 8.3$ or larger), location, and fault models of the Jogan earthquake have been proposed, and the recurrence interval of similar earthquakes and tsunamis was estimated to be somewhere between 500 and 1000 years (Minoura et al., 2001; Sawai et al., 2012).

4.19.3.2.4 Tsunami from the Storegga Slide

The Storegga Slide in the Norwegian Sea is the largest submarine landslide ever mapped. The Storegga Slide is estimated to have taken place three times, once 30–50 thousand years ago (ka) and twice 6–8 ka, with a total volume of 5.6×10^3 km^3 (Bugge et al., 1988). The gigantic submarine landslide more than likely caused a tsunami. Tsunami deposits inferred from this slide have been found on the coasts of Scotland, Iceland, and Norway (Bondevik et al., 2005; Smith et al., 2004).

4.19.3.2.5 Tsunami near the Cretaceous–Tertiary boundary

A global-scale impact tsunami is believed to have occurred at the Cretaceous–Tertiary (K–T, currently known as Cretaceous–Paleogene or K–Pg) boundary about 66 Ma (Schulte et al., 2010). An asteroid of ~10 km diameter impacted the edge of the present Yucatan Peninsula, Mexico, forming the Chicxulub crater (~180 km diameter), and caused catastrophic environmental effects and mass extinctions including the demise of dinosaurs. Because the impact crater was located submarine at the time, the impact must have generated tsunamis. The tsunami deposit attributed to the K–T impact in Texas suggested that the tsunami height was more than 50 m (Bourgeois et al., 1988) and a submarine tsunami deposit found in Cuba was as thick as 100 m (Goto et al., 2008). However, there has been criticism about the interpretation of the tsunami's origin (Keller, 2011).

4.19.4 Tsunami Observations

4.19.4.1 Instrumental Measurements

4.19.4.1.1 Coastal tsunami measurements

Coastal sea-level heights have been measured for more than a century by tide gauges at ports and harbors throughout the world. The classic tide gauge is a mechanical type set in a tide or stilling well. Sea-level change is detected by a float in the well. Because the typical period of ocean tides is several hours or longer, sea-level changes inside the well can be assumed to be the same as that outside the well. For tsunamis with shorter periods (< 5 min), however, the system does not always follow the actual sea-level change (Satake et al., 1988).

More recent gauges include pressure, acoustic, and radar measurements (Intergovernmental Oceanographic Commission, 2006). Pneumatic or bubbler tide gauges use air or gas, and its pressure is measured to estimate the water pressure. A pressure sensor gauge directly measures the water pressure. The water pressure is then converted to the sea-level height. An acoustic gauge sends ultrasonic waves that are reflected at sea surface, and the sea-level changes can be measured from the travel times of the acoustic waves (**Figure 7**). A radar gauge employs the same principle but is less sensitive to temperature than an acoustic gauge and consumes more power for continuous measurements.

The Global Sea Level Observation System, an international program, was formed to establish a high-quality global sea-level network for climate change, oceanography, and coastal research (Intergovernmental Oceanographic Commission, 2006). It consists of approximately 300 tide gauge stations whose data are available in real time through the Internet. Because the main purpose of sea-level measurements is to monitor ocean tides or long-term sea-level changes, the typical sampling interval is not always short enough for tsunami monitoring. For tsunami measurements, a high sampling rate of at least 1 min is required.

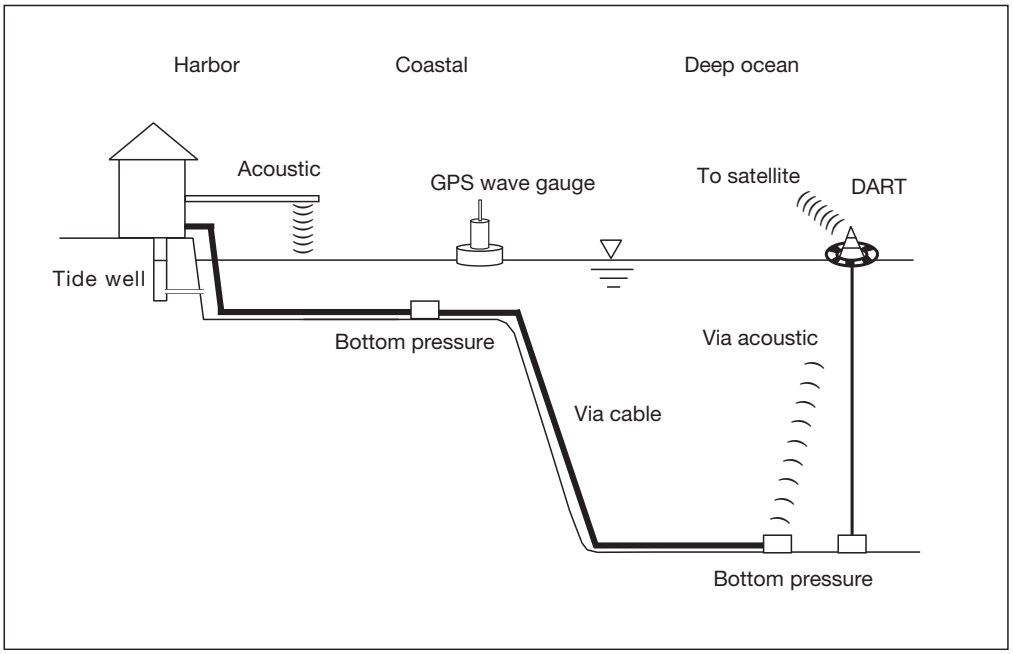


Figure 7 Various types of instruments used to measure tsunamis.

4.19.4.1.2 Offshore tsunami measurements

As a tsunami becomes slow but large and complex along coasts, it is ideal to detect its arrival offshore. Offshore wave gauges, which use ultrasonic waves to measure water height and period, have been developed. Around Japan, approximately 60 wave gauges are installed outside harbors at water depths of tens of meters. In addition, 15 GPS wave gauges have been installed typically 10 km offshore at a water depth of about 100–200 m (Kato et al., 2000). These gauges use the Real-Time Kinematic GPS technique to estimate the location and altitude of the water surface (**Figure 7**). The sampling interval is 1 s with an accuracy of a few centimeters. Because it measures water surface movements, it is free from seismic waves, unlike bottom pressure gauges, and can also be used to monitor wave heights during storms.

4.19.4.1.3 Deep-ocean measurements

In the deep ocean, tsunami waveforms are expected to be simpler (i.e., free from coastal topographic effects), although their amplitudes are smaller. Deep-ocean measurements of tsunamis have been made by using bottom pressure gauges for the early detection and warnings of a tsunami. A quartz crystal sensor is typically used as a pressure sensor. The sensor provides an accuracy of 1 cm in sea level at a 6000 m depth. Because the pressure transducer is very sensitive to temperature, measurements and corrections for temperature are essential. In addition, because the sensor is located at the ocean bottom, it also records ground motion and works as an ocean-bottom seismograph. While the frequencies are often different between seismic and tsunami waves, it is necessary to separate them.

The Deep-ocean Assessment and Reporting of Tsunamis (DART), or simply tsunameter, records water levels using bottom pressure gauges and sends signals to a surface buoy via acoustic telemetry in the ocean and then via satellites to a land station in real time (González et al., 2005a).

Around Japan, more than ten bottom pressure gauges are installed on six submarine cables (Hirata et al., 2002; Iwasaki et al., 1997; Okada, 1995). Tsunamis from the 2004 earthquakes off the Kii Peninsula ($M_w = 7.2$ and 7.4), relatively small tsunamigenic earthquakes, were recorded by a GPS wave gauge and eight bottom pressure gauges, as well as coastal tide gauges located south of Honshu and Shikoku (Satake et al., 2005). The 2011 Tohoku tsunami was also recorded by many bottom pressure gauges (**Section 4.19.2.2**).

Other types of sensors may also be used for tsunami detection. A seafloor geomagnetic observatory in the Northwest Pacific detected clear electromagnetic (EM) variations associated with a tsunami passage from the 2006 and 2007 Kuril earthquakes ($M_w = 8.3$ and 8.1) (Toh et al., 2011). Unlike pressure gauges, EM sensors provide vector measurements from which particle motions of water can be estimated.

4.19.4.1.4 Satellite observations

Four satellites, which were monitoring sea-surface heights (SSH; **Chapter 3.05**), captured the tsunami propagation across the Indian Ocean on 26 December 2004 (**Figure 3**): Jason-1 and TOPEX/Poseidon about 2 h after the earthquake, Envisat 3 h 15 min after the earthquake, and the US Navy's GEOSAT follow on at 7 h (Gower, 2005; Smith et al., 2005). The maximum SSH of the tsunami was about 0.8 m in the Indian Ocean. The accuracy of the satellite altimetry is better than 5 cm, and its spatial resolution is approximately 15 arc minutes (~27 km). A numerical simulation from the 2004 source model was able to reproduce the SSHs (Hirata et al., 2006; Smith et al., 2005) (**Section 4.19.2.1**).

4.19.4.1.5 Definition of amplitude and wave heights

A tsunami's *amplitude* is measured from a zero value to a peak, either the crest (positive peak) or the trough (negative peak). The *wave height*, also referred to as the range or double amplitude, is measured from trough to crest. Hence, the *tsunami height* is measured from trough to crest by tide gauges, whereas the *tsunami amplitude* is measured from the estimated tidal level (Figure 8). As in seismic waves, the *initial motion*, either positive or negative, carries information on the tsunami source. For example, the initial motions of the 2004 tsunami in the Indian Ocean were negative to the east of the source (e.g., Thailand) but positive to the west (e.g., Maldives), as shown in Figure 8.

4.19.4.2 Field Survey to Measure Tsunami Heights

4.19.4.2.1 Posttsunami surveys

Posttsunami field surveys are often conducted after the occurrence of large tsunamis. Usually, experts from different backgrounds, such as coastal engineering, geology, or seismology, and from different countries form an international tsunami survey team and visit the affected coasts to document the effects of the tsunami. An essential element of these surveys is to measure the tsunami heights along the coast (Intergovernmental Oceanographic Commission, 2013). The heights, defined in the succeeding text, are measured from various traces, such as damage to buildings or other structures, watermarks left on walls, vegetative markers that include trim lines and salt-water discoloration or die-off, and lines of debris deposited on beaches. When no physical evidence is found, the tsunami heights are estimated based on eyewitness accounts but considered as less reliable. Locations are usually identified by portable GPS measurements, and the heights are measured by traditional leveling or laser surveying. Interviews with local people also include the degree of earthquake ground shaking, tsunami arrival times, the number of waves with periods, and the duration. Posttsunami surveys also include descriptions and documentations of building damage and tsunami deposits. Synolakis and Okal (2005) reviewed and summarized 15 posttsunami surveys made by international tsunami survey teams during the 11-year period between 1992 and 2002.

4.19.4.2.2 Definitions of inundation and run-up heights

The height of a tsunami on land is called *flow depth* when measured from the ground level, and the *inundation height* is measured relative to the sea level at the time of the tsunami's arrival. The height at maximum inundation, where the velocity and kinetic energy of a tsunami become zero, is called *run-up height*. Because the tide levels at the time of measurement and at the arrival of a tsunami are usually different, corrections are required for the survey data. Unless the sea surface of the inundated tsunami is flat, the run-up height is not necessarily the same as the inundation height or water depth at the shore. The horizontal distance of water penetration measured from the coast is called the *inundation distance* (Figure 9).

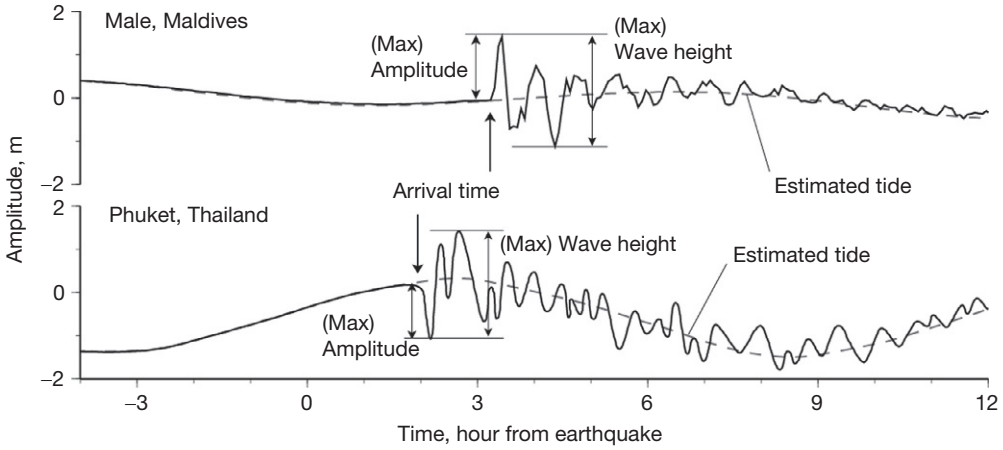


Figure 8 Tsunami waveforms from the 2004 Sumatra–Andaman earthquake recorded at Male, Maldives (top), and Phuket, Thailand (bottom). See Figure 3 for the locations.

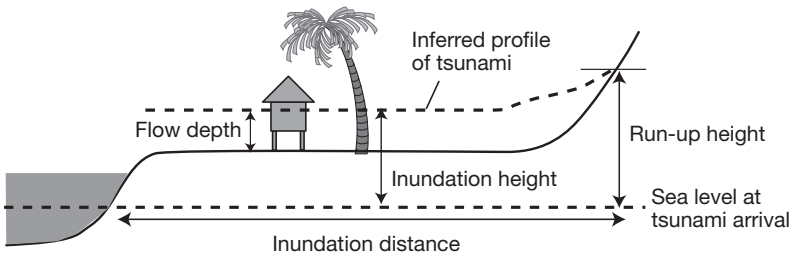


Figure 9 Definitions of tsunami heights and inundation distances on land.

4.19.4.3 Study of Historical Documents

Tsunamis and their effects documented in historical records are also used to study old earthquakes and tsunamis for which no instrumental data are available. Some countries in Europe, South America, and Asia have a long history of written records that provide accurate dates for past earthquakes and tsunamis. For example, tsunami heights from the 1755 Lisbon earthquake have been estimated from historical documents as 6 m in Lisbon, more than 15 m in southwest Portugal, and more than 10 m in Spain and Morocco (Baptista et al., 1998).

Estimation of tsunami heights from historical damage descriptions requires numerous assumptions. For example, the Japanese descriptions of the January 1700 tsunami, generated by a giant earthquake in the Cascadia subduction zone (**Section 4.19.3.2.1**), included flooding in farmed fields and a government warehouse and damage to coastal houses (Atwater et al., 2005). To convert these reported effects into estimated heights, Satake et al. (2003) considered various factors, such as the relationship between tsunami damage and heights, sea-level change in the last 300 years, and similarity with the 1960 Chilean tsunami. Japanese tsunami data indicate that the amount of structural damage is small when the tsunami height is less than 2 m but becomes significant when tsunami heights are greater than 4 m (Hatori, 1984). On the basis of the 2004 tsunami damage data for Banda Aceh, a tsunami fragility curve, which indicates the damage probability as a function of inundation depth, was constructed (Koshimura et al., 2009). The curve shows that the damage probability is about 20% for 2 m and nearly 80% for a 4 m inundation depth.

4.19.4.4 Geologic Methods

Geologists have studied the deposits of prehistoric tsunamis (e.g., Bourgeois, 2009; Dawson and Shi, 2000; Goff et al., 2012). In Europe, inferred tsunami deposits from the Storegga Slides in the North Sea have been found and studied in Scotland, Norway, and Iceland (Dawson et al., 1988). In the United States, tsunami deposits from prehistoric great earthquakes along the Cascadia subduction zone have been studied (Atwater et al., 2005). In Australia, geologic evidence of late Pleistocene tsunamis has been found (Bryant et al., 1996). In Japan, tsunami deposits or boulders in various environments, such as lagoons, marshes, and beaches, have been studied for modern, historic, and prehistoric tsunamis (Goto et al., 2010; Minoura and Nakaya, 1991; Nanayama et al., 2003). Tsunami geology can help identify tsunami hazards that would otherwise come as a surprise, and it can help quantify the earthquake and tsunami hazards in terms of sizes and recurrence intervals (**Chapters 4.21** and **4.22**).

Tsunami deposits are sedimentologically classified as event deposits, resulting from relatively short and rare episodes compared to the usual geologic depositional processes. Events include floods, storm surges due to a hurricane or cyclone, and tsunamis. Because tsunamis produce long waves and storms produce short waves (see **Section 4.19.6.4**), some differences in their sedimentologic characteristics are expected. In general, tsunami deposits are transported as a suspended load by multiple waves, deposited rapidly in sheets, and show graded bedding, whereas storm deposits are transported by bed load and continuously deposited forming cross-bedding (Morton et al., 2007; Nanayama et al., 2000; Tuttle et al., 2004). Once the deposit is attributed to its tsunami origin, the source and date need to be estimated for paleoseismological purposes. If a tsunami is from a near-field earthquake, other phenomena such as coseismic uplift/subsidence or liquefaction may also be recorded as a geologic description (Nelson et al., 1996; Shimazaki et al., 2011; **Chapter 4.21**). Studies of deposits from modern tsunamis are important; hence, careful descriptions have been made for recent tsunamis (e.g., see **Section 4.19.3**). The previously mentioned case studies indicate that the characteristics of tsunami deposits are highly variable and depend on coastal topography and geology.

4.19.4.5 Tsunami Magnitude Scales

4.19.4.5.1 Imamura–Iida scale (m)

The traditional tsunami magnitude scale is the Imamura–Iida scale, m. Although the original definition is descriptive, the value of m can be approximately computed as

$$m = \log_2 h \qquad [1]$$

where h is the maximum run-up height in meters (Iida et al., 1967). This scale is similar to the earthquake intensity scale and is especially convenient for old tsunamis of which no instrumental records exist.

Hatori (1979) extended Imamura–Iida's m scale to include far-field tsunami data as

$$m = 3 + \log\left\{(h/0.5)(R/1000)^{1/2}\right\}/\log\sqrt{5} \qquad [2]$$

where h is in meters and R is the distance from the tsunami source to the coast in kilometers. This indicates that the tsunami height decays with distance as $1/\sqrt{R}$, which is theoretically predicted for nondispersive waves traveling a long distance (Comer, 1980).

4.19.4.5.2 Tsunami intensity scale (I)

Soloviev (1970) pointed out that Imamura–Iida's m scale is more like an earthquake intensity scale rather than a magnitude scale. He also distinguished the maximum tsunami height h and the mean tsunami height H_{av}. He then defined the tsunami intensity I as

$$I = \log_2\left(\sqrt{2}H_{av}\right) \qquad [3]$$

Comparison of eqns [1] and [3] suggests that the mean tsunami height H_{av} is given as $1/\sqrt{2}$ times the maximum height h. The maximum intensity I on the coast that is closest to the tsunami source is used to quantify the tsunami source.

4.19.4.5.3 Tsunami magnitude (M_t)

The tsunami magnitude scale, M_t, was defined and used for numerous earthquakes (Abe, 1979, 1981). The definition of M_t for a transpacific tsunami is

$$M_t = \log H + C + 9.1 \qquad [4]$$

and for a regional tsunami ($100\ \text{km} < \Delta < 3500\ \text{km}$) is

$$M_t = \log H + \log \Delta + 5.8 \qquad [5]$$

where H is the maximum amplitude recorded by the tide gauges in meters, C is the distance factor depending on a combination of the source and the observation points, and Δ is the nautical distance in kilometers. The aforementioned formulas were calibrated with the moment magnitude scale, M_w, of the earthquakes. Equation [5] indicates that for the same M_t, the tsunami amplitude H decays with $1/\Delta$, which is different from the assumption used in Imamura–Iida's scale m, and may be valid only for regional distances.

4.19.4.6 Tsunami Catalogs

The location, size, run-up heights, and damage from past tsunamis have been compiled by several researchers and published as catalogs. For the Pacific tsunamis, the Iida catalog (Iida et al., 1967) contains numerous tsunami height data that date back as far as AD 173. Russian catalogs (Soloviev and Go, 1974a,b; Soloviev et al., 1992) include Pacific tsunami data (except for Russia) with some tide gauge records, tabulated tsunami heights, and descriptions of each tsunami. Regional catalogs including tide gauge records were also published for Alaska (Lander, 1996), the West Coast of the United States (Lander et al., 1993), Japan (Iida, 1984; Watanabe, 1998), and the Indian Ocean (Berninghausen, 1966; Murty and Rafiq, 1991).

Digital tsunami databases have been compiled, such as the Historical Tsunami Databases at the Institute of Computational Mathematics and Mathematical Geophysics, Novosibirsk, Russia (Gusiakov, 2005), and at the National Geophysical Data Center (NGDC), NOAA, the United States (Dunbar et al., 2008). The NOAA Tsunami Database consists of tsunami source events (more than 2000 events) and tsunami run-up data (more than 20 000 entries). The NOAA database also contains tsunami deposit data (Dunbar and McCullough, 2012). Tsunami heights and traces recorded or estimated from historical documents in Japan were compiled into the Japan Tsunami Trace Database by Tohoku University and the Japan Nuclear Energy Safety Organization. A unique aspect of this database is that the reliability of each entry was examined by a committee of experts.

4.19.5 Tsunami Generation

In this section, the generation of tsunamis by earthquakes, landslides, and volcanic eruptions is reviewed.

4.19.5.1 Earthquake Tsunamis

4.19.5.1.1 Seafloor and water surface displacements

Analytic solutions for linear gravity waves generated from several types of sources have been obtained by various researchers. Kajiura (1963) generalized the solutions by introducing the tsunami Green's functions. He showed that the initial conditions given at the seafloor and at the water surface differ by a factor of $1/\cosh kd$, where k is the wave number of the seafloor displacement and d is water depth. For large wavelengths of $\lambda \gg 2\pi d$, this factor becomes 1, since $\cosh kd = \cosh(2\pi d/\lambda) \approx 1$. Therefore, if the wavelength of the seafloor displacement is much larger than the water depth, it can be assumed that the water surface displacement is the same as the seafloor displacement.

Kajiura (1970) examined the energy exchange between the solid Earth and ocean water for tsunamis generated by seafloor displacements with a finite rise time. He showed that the efficiency of tsunami generation, defined as the ratio of dynamic tsunami energy to static energy, becomes nearly one if the duration is short compared to the time required for a tsunami to travel through the source area. For an earthquake with $M \sim 8$ at a water depth of 4000 m, the source size is about 100 km and the tsunami travel time over that distance is about 8 min. If the duration is much shorter, for example, 1 min or less, the efficiency is more than 0.9. This means that the seafloor displacement can be considered to be instantaneous. If the duration is too short, shorter than $2\sqrt{d/g}$, a significant part of the energy flux goes into an acoustic wave rather than tsunami wave. For the aforementioned case, this limit is 45 s.

4.19.5.1.2 Fault parameters and seafloor displacement

Seafloor displacements due to faulting can be calculated using the elastic theory of dislocation (**Chapter 4.10**). The displacement in an infinite homogeneous medium due to a dislocation across a fault is given by Volterra's theorem (Steketee, 1958), and the explicit analytic formulas for half-space have also been given (Mansinha and Smylie, 1971; Okada, 1985). Usually, only the vertical components of seafloor displacements are considered during the analysis of a tsunami's generation. When the tsunami source is on a steep seafloor slope and the horizontal displacement is large, the vertical displacement of water due to the horizontal displacement of the slope must be also considered (Tanioka and Satake, 1996b). For the 2011 Tohoku tsunami, the large horizontal displacement of seafloor slope was responsible for 20–40% of the observed tsunami amplitudes (Satake et al., 2013a).

The fault parameters needed to compute seafloor displacement include the fault location, geometry (strike, dip, and rake angles), size (length L and width W), and average slip, $\bar{u}$ (or the strike-slip and dip-slip components). These parameters can be estimated using a seismic wave analysis or aftershock distributions. The seismic moment M_0 is given as

$$M_0 = \mu \bar{u} S = \mu \bar{u} L W \qquad [6]$$

where S is the fault area ($L \times W$) and μ is the rigidity around the fault. The moment magnitude is defined as

$$M_w = \frac{2}{3} \log M_0 - 10.7 \qquad [7]$$

where M_0 is given in dyne cm (10^{-7} Nm) (Hanks and Kanamori, 1979). The effects of the fault parameters on tsunami generation were examined by Yamashita and Sato (1974) and Geist (1998).

4.19.5.1.3 Estimation of earthquake source parameters from tsunamis

Tsunami data can be used to study earthquake source processes in a way similar to that in which seismic waves are used (**Chapter 4.09**). Abe (1973) analyzed the tsunami arrival times and first-motion data from the 1968 Tokachi-oki earthquake. By drawing inverse refraction diagrams from each station (**Section 4.19.6.2**), he estimated the tsunami's source

area, which agrees well with the aftershock area. From the first motion of the tsunami waves, the initial water surface disturbance was estimated as uplift at the southeastern edge and subsidence at the northwestern edge. This pattern is similar to vertical seafloor displacements due to faulting, which are independently estimated from seismological analyses.

Numerical simulation of a tsunami has been carried out for many fault models around Japan (Aida, 1978). The best fault models were obtained by conducting trial and error comparisons of observed and computed tsunami waveforms or coastal tsunami heights. Tsunami waveforms recorded by tide gauges were then inverted to estimate the slip distribution on the fault (Satake, 1989). In this inversion method, the fault plane is first divided into several subfaults, and the seafloor displacement is computed for each subfault with a unit amount of slip. Using this as an initial condition, tsunami waveforms are numerically computed via actual bathymetry. The observed tsunami waveforms are expressed as a superposition of waveforms computed for each subfault as

$$A_{ij}(t) \cdot x_j = b_i(t) \quad [8]$$

where $A_{ij}(t)$ is the computed waveform as a function of time t, or the Green's function, at the ith station from the jth subfault; x_j is the amount of slip on the jth subfault; and b_i (t) is the observed tsunami waveform at the ith station. The slip x_j on each subfault can be estimated by a least-square inversion of the aforementioned set of equations. The tsunami waveform inversion has been applied to many large and great earthquakes in the Alaska–Aleutians (Johnson, 1999), northern Japan (Satake, 1989; Tanioka et al., 1995), and the Nankai Trough (Satake, 1993; Tanioka and Satake, 2001a,b) and to the 2004 Sumatra–Andaman earthquake (Fujii and Satake, 2007). For the 2011 Tohoku earthquake, because high-quality and high-sampling offshore tsunami waveforms were available, the temporal change and spatial distribution of the slips on subfaults were estimated (Satake et al., 2013a). These results show that the slip distributions are not uniform but are concentrated in small regions, which correspond to asperities (**Figures 3** and **5**).

The tsunami inversion method has been expanded to adopt an adjoint method for nonlinear waves (Pires and Miranda, 2001) in order to include coastal tsunami heights (Piatanesi et al., 1996) and to enable the real-time assimilation of DART data for the purpose of providing tsunami warnings (Titov et al., 2005b) as described in **Section 4.19.7.1**.

4.19.5.1.4 Tsunamigenic and tsunami earthquakes

An earthquake that generates a tsunami is called a tsunamigenic earthquake. A tsunami earthquake, on the contrary, is defined as an earthquake that excites much larger tsunamis than expected from its seismic waves (Kanamori, 1972). This discrepancy can be quantified in terms of the surface wave magnitude (M_s) and tsunami magnitude (M_t) scales; when M_t is larger than M_s by more than 0.5 units, a tsunami is classified as a tsunami earthquake (Abe, 1989). Newman and Okal (1998) introduced a new parameter called the slowness parameter, $\Theta = \log_{10}(E^E/M_0)$, which is the ratio of radiated seismic energy E^E to seismic moment M_0. They showed that Θ is about -5 for regular tsunamigenic earthquakes and about -6 or smaller for tsunami earthquakes. They proposed to use this parameter for the discrimination of tsunami earthquakes.

Several proposals have been made for the generating mechanisms of tsunami earthquakes. Slow and long rupture processes reconcile the large discrepancy between seismic and tsunami waves (Kanamori, 1972). High-angle secondary faulting in the accretionary wedge can explain the relatively large tsunami excitation (Fukao, 1979). For a given slip, the seismic moment is smaller in the shallower portion of a subduction zone because of the smaller rigidity, but the tsunami becomes larger (Okal, 1988). Tsunami heights may not be anomalously large if compared to M_w rather than M_s (Pelayo and Wiens, 1992); the large discrepancy between the tsunami and M_s is due to the saturation of M_s at around 7.3 for those earthquakes that occurred in the accretionary wedge.

Recent tsunami waveform modeling showed that the fault parameters and locations of the tsunami earthquakes share a common feature: narrow and shallow faulting near the trench axis (Satake and Tanioka, 1999). The 2 September 1992 Nicaragua earthquake (M_s=7.2 and M_t=7.9) was the first tsunami earthquake recorded by modern broadband seismic instruments. Seismological studies showed that the duration was very long for its size, which was approximately 100 s (Kanamori and Kikuchi, 1993). Comparison of numerically computed tsunami waveforms with tide gauge records (Satake, 1994) showed that a narrow (~40 km) and shallow (extending only to the upper 10 km of the ocean bottom) fault was responsible for the tsunami generation.

Other tsunami earthquakes occurred along the Sunda Trench south of Java Island, Indonesia, in 1994 and 2006 (Fujii and Satake, 2006; Polet and Kanamori, 2000). While only typical interplate or tsunamigenic earthquakes had been recorded off the coast of Sumatra Island, the 2010 Metawai earthquake (M_s=7.3 and M_t=8.3) occurred at the trench side of the source region of the 2007 Bengkulu earthquake (Fujii and Satake, 2008) and is considered to be a tsunami earthquake (Satake et al., 2013b) (**Figure 10**). Among the tsunami earthquakes, the 1896 Sanriku earthquake (M_s=7.2 and M_t=8.6) and the 1946 Aleutian earthquake (M_s=7.3 and M_t=9.3) were very anomalous (Kanamori, 1972; Kanamori and Kikuchi, 1993). Tsunami models showed a similar fault parameter for the tsunami earthquakes, that is, a narrow fault was located near the trench axis (Johnson and Satake, 1997; Tanioka and Satake, 1996b; Tanioka and Seno, 2001a,b). Tanioka et al. (1997) showed that the 1896 tsunami earthquake occurred in a region where the ocean-bottom topography was rough and characterized by well-developed horst and graben structures. Polet and Kanamori (2000) extended this model to global subduction zones, based on the examination of the source spectra of large ($M > 7$) earthquakes in the 1990s. Okal and Newman (2001), however, failed to find any correlation between the slowness parameter Θ and the depth, focal mechanism, or seismic moment for Central America, Java, and Peru, suggesting an absence of the regional signal for the occurrence of tsunami earthquakes. Bilek and Lay (2002), from the analysis of the source time functions of tsunamigenic earthquakes and tsunami earthquakes, found that shallower earthquakes have longer duration and claimed that the tsunami earthquakes are simply larger earthquakes at a very shallow depth. Seno (2002) proposed that an increase in fluid

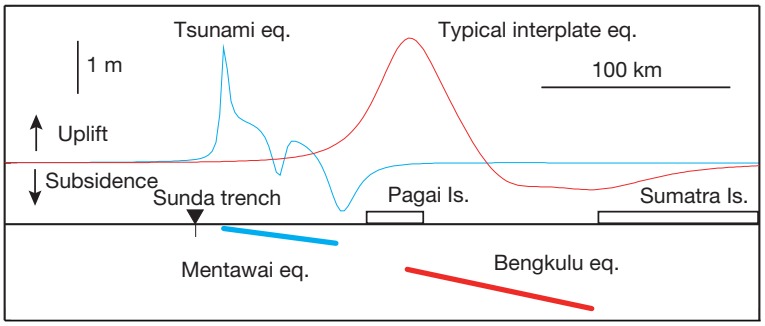


Figure 10 Schematic figure showing the vertical seafloor displacement (top) due to faulting (bottom) for the 2007 Bengkulu earthquake (typical interplate earthquake, red) and the 2010 Mentawai earthquake (tsunami earthquake, blue), based on the results of Fujii and Satake (2008) and Satake et al. (2013b).

pressure causes a temporal change in the frictional properties near the trench axis and is responsible for slow ruptures and therefore concluded that tsunami earthquakes are transient phenomena.

4.19.5.2 Tsunamis Caused by Landslides

Tsunamis caused by submarine landslides have been a hot research topic since the 1990s. Abnormally large and concentrated tsunami heights from the 1992 Flores Island (Indonesia) and 1998 Sissano (Papua New Guinea) earthquakes and the nonseismic 1994 Skagway (Alaska) tsunami provide case studies for such occurrences. Marine surveys revealed detailed bathymetric features of landslides or slumps. Physical experiments in tanks and numerical computations of landslide-generated tsunamis have been developed. Such multidisciplinary studies have been summarized in several special issues (Bardet et al., 2003; Keating et al., 2000; Tappin, 2004).

4.19.5.2.1 Examples

Posttsunami field surveys after large tsunamigenic earthquakes sometime show abnormally large tsunami heights concentrated along small segments of coasts. While offshore bathymetry or coastal topography (such as V-shaped bays) may provide such results for incoming tsunamis, local large tsunamis may be caused by a local landslide, either subaerial or submarine, that was triggered by ground shaking. Large water waves were caused by a rockslide in Lituya Bay at the time of the 1958 earthquake (Section 4.19.3.1.4). Concurrent with the 1964 Alaska earthquake (M_w=9.2), in addition to Pacific-wide tsunamis, local tsunamis from landslides were reported at more than 20 locations; the largest run-up height was 50 m (Lander, 1996; Plafker et al., 1969). The tsunami heights of the 1992 Flore Island earthquake were mostly 4–5 m along most of the northern coast, but 26 m at Riangkroko (Yeh et al., 1993). The tsunami heights from the 1998 Papua New Guinea tsunami, described in Section 4.19.3.1.1, show a local peak of 15 m (Figure 11).

Larger-scale submarine slumping and turbidity currents were triggered by the 1929 Grand Banks earthquake (Fine et al., 2005a; Heezen and Ewing, 1952). With a maximum tsunami run-up height of 13 m and 28 casualties, it was the worst tsunami disaster in Canada. The turbidity current also caused damage to the transatlantic submarine cables at 12 locations over a total distance of 600 km.

The 1994 Skagway tsunami is an example of a nonseismic landslide tsunami (Kulikov et al., 1996). A submarine landslide, without a seismic trigger, occurred in Skagway Harbor in Alaska and generated a local tsunami with heights of 5–10 m. The tide gauge records show a dispersive character (Section 4.19.6.6) due to the small size of the source.

4.19.5.2.2 Characteristics and discrimination of landslide tsunamis

The scale length for submarine mass failure is very different than that of submarine faulting. While vertical movement can be hundreds of meters, the horizontal scale is only, for the most part, a few to several kilometers for submarine mass failure. Seafloor displacement due to faulting, on the contrary, has at most several meters of vertical displacement, but the spatial extent can be several hundreds of kilometers. As will be discussed in the next section, an earthquake source almost always satisfies the long-wave (shallow-water) conditions, while submarine landslides may not. Hence, the water waves from submarine landslides show a dispersive characteristic and attenuate quickly without reaching a long distance.

Okal and Synolakis (2004) fit the run-up height distribution $\zeta(y)$ along a beach (in the y direction) by using the following function:

$$\zeta(y) = \frac{b}{[(y-c)/a]^2 + 1} \qquad [9]$$

where parameters a, b, and c are estimated from the observed data (Figure 11). Okal and Synolakis (2004) showed that $I_2 = b/a$, which is the aspect ratio of the maximum run-up heights to the extent of the run-up along the beach, can be used as a discriminant for the nature of the tsunami source and for the dislocation (earthquake fault) or landslide source. Seismic dislocation sources have I_2 smaller than 10^{-4}, whereas a landslide source has a larger value.

4.19.5.2.3 Modeling landslide tsunamis

Submarine landslide sources of tsunamis have been also modeled in many ways. These can be grouped into four categories: viscous fluid model, rigid-body or block model, initial static water surface profile, and moving kinematic water surface profile.

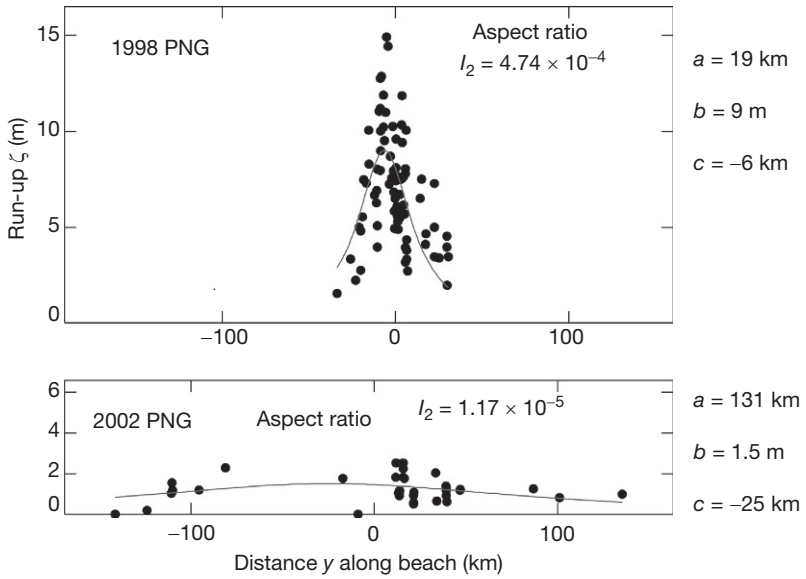


Figure 11 Comparison of tsunami run-up heights for the 1998 and 2002 Papua New Guinea tsunamis. The estimated parameters of eqn [9] are also shown. Modified from Okal EA and Synolakis CE (2004) Source discriminants for near-field tsunamis. *Geophysical Journal International* 158: 899–912.

In the viscous fluid approach, a landslide is treated as a fluid and its evolution is calculated simultaneously with tsunami generation. Jiang and Leblond (1994) treated a landslide as incompressible viscous flow and sea water as inviscid flow and formulated the multiphase fluid dynamics. Rzadkiewicz et al. (1997) modeled landslides and water as a viscoplastic fluid and dispersive ideal fluid, respectively, and compared the results with analytic solutions and physical experiments. Fine et al. (2003) modeled the subaerial and submarine slides as viscous flows and compared the results with a rigid-body model. They found that the viscous flow and rigid-body motion produce different waves for the submarine landslides. Heinrich et al. (2001), Titov and Gonzalez (2001), and Imamura and Hashi (2003) applied similar methods to the submarine landslide of the 1998 Papua New Guinea tsunami. Kawamata et al. (2005) modeled the submarine landslide of the 1741 Oshima-Oshima tsunami with a two-layer model.

In the rigid-body approach, the temporal change of water height due to the progression of submarine landslides is prescribed and used to compute tsunami propagation. Harbitz (1992) assumed the Storegga slide to be box-shaped. Ward (2001) used a chain of rigid boxes to model the Storegga and other submarine slides. Tinti et al. (1997) introduced a block model for landslides and simulated the motion of a landslide. Tinti et al. (2005) applied this method for the 2002 tsunami that was generated near the Stromboli volcano.

In the static approach, an initial profile caused by a submarine landslide is assumed. Based on the results of physical and numerical experiments, Grilli and Watts (2005) introduced a parametric initial profile. Watts et al. (2003) applied these initial profiles for several landslide-generated tsunamis including the 1946 Aleutian, 1994 Skagway, and 1998 Papua New Guinea tsunamis. Synolakis et al. (2002) and Satake and Tanioka (2003) used the same approach to model the 1998 Papua New Guinea tsunami. This approach, however, does not consider the landslide propagation and ignores the kinematic energy from the submarine landslide.

The last approach assumes a moving water surface profile as the initial motion of tsunami propagation. Todorovska et al. (2003) introduced five kinematic models to represent the submarine landslides and slumps. Satake et al. (2002) introduced a simple kinematic landslide model on the basis of a change in bathymetry and applied it to the Hawaiian landslides. These studies found that slide velocity is one of the important parameters in controlling the amplitudes of a tsunami. The same procedure was also applied to the 1741 Oshima-Oshima slide (Satake, 2007).

For water waves generated from subaerial landslides, various studies using theoretical, experimental, numerical, and field approaches are summarized in Slingerland and Voight (1979).

4.19.5.3 Volcanic Tsunamis

Tsunamis are one of the most important hazards associated with volcanic activity (**Chapter 4.14**). About a quarter of human casualties directly associated with volcanic eruptions are due to tsunamis (Tilling, 1989). From case studies of 106 volcanic tsunamis, Latter (1981) classified them into ten groups. Volcanic tsunamis can be roughly categorized into five groups according to their origin: (1) earthquakes; (2) landslides and avalanches of cold rock and lahars; (3) flow of volcanic ejecta such as pyroclastic flow, base surges, hot rock avalanches, or lava; (4) submarine explosions; and (5) caldera collapses or subsidence. Each of these categories is briefly introduced in the succeeding text with examples.

Volcanic earthquake tsunamis are generated by the same mechanism, seafloor displacement due to faulting, as described in the previous section, except that the earthquakes are of volcanic origin. Examples of volcanic earthquakes include

the 1914 Sakurajima (Japan), 1930 Stromboli (Italy), and 1933 Severgin (Kuril Islands) tsunamis.

A typical example of a volcanic-landslide tsunami is one that occurred in Shimabara, Japan. In 1792, Mayuyama, a small mountain near the Unzen volcano, collapsed and the resulting landslide entered Ariake-kai, a shallow bay with a typical depth of 10–20 m, which generated a tsunami that killed ~15 000 people. Because the horizontal extent of the slide was about 3000 m, much longer than the water depth, this satisfied the long-wave approximation. Thus, the tsunami heights were large (10–20 m) on the other side of the bay, which had significant damage (~5000 casualties).

Emplacements of volcanic ejecta in the water also cause tsunamis. Examples of pyroclastic flow include St. Vincent and Mt. Pelee in the Lesser Antilles in 1902 and St. Augustine (Alaska) in 1883. Maeno and Imamura (2011) considered the 1883 Krakatoa tsunami to be caused by pyroclastic flow. Other types of ejecta include base surge from the Taal volcano in the Philippines; a hot rock avalanche in Stromboli, Italy; and a lava avalanche from the Matavanu (Samoa) volcano in 1906–07.

An example of a tsunami that was caused from a submarine explosion is the eruption of the Bayonnaise Rocks (located along the Izu–Bonin Islands about 300 km south of Honshu, Japan) in 1952. The tsunami was recorded at Hachijo Island, about 130 km north of the source, and had a very-short-period (~1 min) wave and a dispersive character. Unoki and Nakano (1953) applied a Cauchy–Poisson wave (deep-water waves from initial water surface disturbances) to the observed tsunami record and estimated the source size (~2 km) and water surface uplift (impulse force).

The final mechanism of a volcanic tsunami is a caldera collapse or subsidence. Large-scale caldera collapses were recorded at the times of the 1883 Krakatoa and the 1883 Ritter Island (Papua New Guinea) eruptions. However, Ward and Day (2003) modeled the Ritter Island tsunami by using a landslide model.

4.19.6 Tsunami Propagation

4.19.6.1 Weak Coupling between the Elastic Earth and Ocean

In the Earth-ocean coupled system, an earthquake source would generate waves both in the elastic Earth and in the ocean (**Chapter 4.16**), and the propagation of a tsunami in the ocean layer would excite waves in the elastic Earth. In contrast, the traditional decoupled approach first calculates the elastic deformation on the ocean bottom generated by an earthquake and then uses this as an initial condition for tsunami wave propagation such that the tsunami wave in the ocean would have no influence on the elastic Earth. The theoretical treatment of a tsunami as a coupled system is similar to that of Rayleigh waves or spheroidal oscillations (**Chapters 1.03** and **1.04**). Ward (1980) and Okal (1988) extended a normal model approach using the spherical coordinate system to solve the generation and propagation of tsunamis, while Comer (1984b) extended the Rayleigh wave approach using a Cartesian coordinate system to solve the tsunami problem. The calculated dispersion curve shows that, except at very low frequencies, the phase velocity is very similar to that of the gravity wave. Both approaches give almost identical solutions due to the small compressibility of water (Comer, 1984a). In other words, the coupling between the elastic Earth and the ocean is extremely weak, which explains why very large earthquakes are necessary to generate tsunamis.

Recently, the effect of coupling with the elastic Earth has been attributed to the large arrival-time differences between the observed and predicted tsunamis across the Pacific Ocean (Tsai et al., 2013; Watada et al., 2013). For recent transpacific tsunamis, such as the 2007 Peru tsunami, the 2010 Chile tsunami, and the 2011 Tohoku tsunami, discrepancies (a few percent) in the travel time between the observed waveforms recorded by offshore tsunami gauges (DART or GPS wave gauges) and the computed waveforms have been observed (Kato et al., 2011; Tang et al., 2012; Wei et al., 2008). The smaller phase velocity of the dispersion curve at the lowest frequency, due to the coupling of elastic Earth and water, is considered to be responsible for these delays.

4.19.6.2 Ray-Theoretical Approach

If the tsunami wavelength is much smaller than the scale of heterogeneity in propagation velocity, or the change in water depth, then the geometrical ray theory of optics can be applied (**Chapter 1.05**). The wave fronts of a propagating tsunami can be drawn on the basis of Huygens' principle. Such a diagram is called a *refraction diagram*. Refraction diagrams can be prepared for major tsunami sources and can be used for tsunami warnings; as soon as the epicenter is known, the tsunami arrival times can be readily estimated. Some examples can be found at the NOAA NGDC Tsunami Database. **Figure 2** shows the refraction diagrams for the 2004 Andaman–Sumatra and 2011 Tohoku earthquakes and the wave fronts for each hour. The 2004 Indian Ocean tsunami was expected to arrive at the Thai and Indian coasts within 2 h and at the African coast within 7 h of the earthquake, while the 2011 Tohoku tsunami was expected to arrive at the North and South American coasts within 8 and 18 h, respectively.

The refraction diagrams can be drawn from a coastal point. If the wave fronts or rays are traced backward from the coastal observation points for the corresponding travel times, then they encompass the tsunami source area (Miyabe, 1934). This type of diagram is called an *inverse refraction diagram* and has been used to estimate the tsunami source area. The tsunami source extent of the 2004 Sumatra–Andaman earthquake was estimated from inverse refraction diagrams (Lay et al., 2005; Neetu et al., 2005) (**Section 4.19.2.1**).

4.19.6.3 Coastal Amplification and the Green's Law

The conservation of potential energy along the rays can be shown as

$$b_0 d_0^{1/2} h_0^2 = b_1 d_1^{1/2} h_1^2 \qquad [10]$$

where d is the water depth, b is the distance between the rays, h is the tsunami amplitude, and the subscripts 0 and 1 indicate two different locations in the ocean. If the tsunami amplitude at location 0 (e.g., deep water) is known, the tsunami amplitude at location 1 after the propagation can be estimated as

$$h_1 = \left(\frac{b_0}{b_1}\right)^{1/2} \left(\frac{d_0}{d_1}\right)^{1/4} h_0 \qquad [11]$$

Equation [11] is known as the Green's law. Here, the ratio b_1/b_0 represents the spreading of rays, which can be graphically obtained from refraction diagrams. Note that this applies only to direct waves and may not be applied to reflected or edge waves.

The Green's law can be used to estimate the shoaling effects, an increase of wave amplitudes at shallow depths. For plane waves approaching the coast, the spreading ratio is one; hence, the amplitude is proportional to one-fourth of the power of water depth change. For example, when the water depth becomes one-tenth, the amplitude becomes 1.8 times larger.

4.19.6.4 Tsunamis as Gravity Waves

In this section, a two-dimensional Cartesian coordinate system is used with the z-axis as vertically upward (its origin is on the undisturbed water level) to consider the propagation of a wave (Figure 12). When gravity is the restoring force, it is called a gravity wave. The Euler's equation of motion can be written as

$$\frac{\partial \mathbf{V}}{\partial t} + \boldsymbol{V} \cdot \nabla \boldsymbol{V} = -\boldsymbol{g} - \frac{1}{\rho} \nabla p \qquad [12]$$

where $\boldsymbol{V}$ is the velocity vector (whose x and z components are u and w, respectively), $\boldsymbol{g}$ is the gravitational acceleration, and ρ and p are water density and pressure, respectively.

The phase velocity c of gravity waves is given as

$$c = \left(\frac{g}{k} \tanh kd\right)^{1/2} = \left(\frac{g\lambda}{2\pi} \tanh \frac{2\pi d}{\lambda}\right)^{1/2} \qquad [13]$$

where k is the wave number, λ is the wavelength, and d is the water depth.

When the horizontal scale of motion, or the wavelength, is much larger than the water depth ($\lambda \gg d$), the vertical acceleration of water is negligible compared to gravity. This means that the horizontal motion of water mass is almost uniform from the ocean bottom to the surface (Figure 12). Such a wave is called a shallow-water wave or a long wave. For tsunamis, typical ocean depth is about 4–5 km, and a large earthquake has a source area of several tens to hundreds of kilometers. Hence, the long-wave approximation is appropriate for most cases.

On the other hand, when the wavelength is comparable to the water depth, eqn [13] becomes $c = \sqrt{\frac{g}{k}} = \sqrt{\frac{g\lambda}{2\pi}}$, that is, velocity depends on wavelength. Such a wave is called a deep-ocean wave or a short wave (Figure 12). The longer wavelength components of the waves propagate faster than the shorter wavelength components; hence, the wave shows dispersion and cannot reach a long distance with the initial profile. Wind waves or landslide-generated waves may be approximated as deep-water waves.

If the amplitude is small compared to the water depth ($h \ll d$), the equation of motion, or conservation of momentum, can be written as

$$\frac{\partial u}{\partial t} = -g \frac{\partial h}{\partial x} \qquad [14]$$

and the equation of continuity, or conservation of mass, can be written as

$$\frac{\partial h}{\partial t} = -\frac{\partial}{\partial x}(\mathrm{d}u) \qquad [15]$$

Such a wave is called a small-amplitude wave or linear long wave. The assumptions are valid for most tsunami propagation paths except for nearshore regions where the nonlinear effects become important.

From eqns [14] and [15], by assuming that water depth d is constant, the wave equation can be obtained as

$$\frac{\partial^2 h}{\partial t^2} = c^2 \frac{\partial^2 h}{\partial x^2} \quad \text{where } c = \sqrt{gd} \qquad [16]$$

in which the phase velocity is determined by only the water depth. Note that the aforementioned phase velocity can also be obtained from eqn [13] if $\lambda \gg d$.

4.19.6.5 Basic Equations for Tsunami Propagation

Including the bottom friction and the Coriolis force, the momentum equations for shallow-water (long) waves can be written for a three-dimensional case as

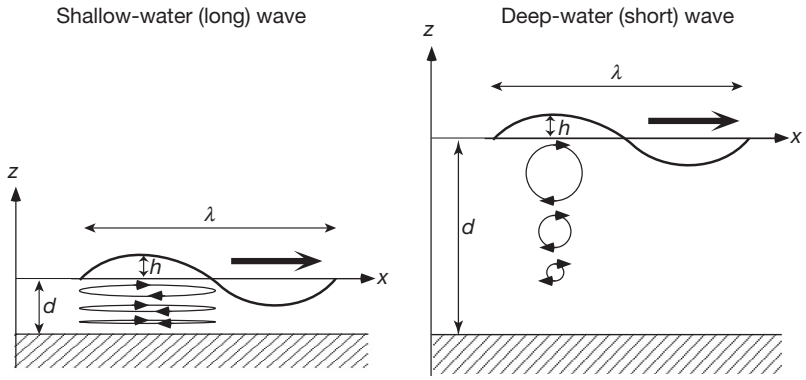


Figure 12 Definition of the coordinate system and the parameters for gravity waves. (Left) Shallow-water-wave, or long-wave, approximation. (Right) Deep-water-wave, or short-wave, approximation.

$$\frac{\partial U}{\partial t}+U\frac{\partial U}{\partial x}+V\frac{\partial U}{\partial y}=-fV-g\frac{\partial h}{\partial x}-C_f\frac{U\sqrt{U^2+V^2}}{d+h}$$

$$\frac{\partial V}{\partial t}+U\frac{\partial V}{\partial x}+V\frac{\partial V}{\partial y}=fU-g\frac{\partial h}{\partial y}-C_f\frac{V\sqrt{U^2+V^2}}{d+h} \quad [17]$$

The mass equation can be written as

$$\frac{\partial h}{\partial t}+\frac{\partial}{\partial x}\{U(h+d)\}+\frac{\partial}{\partial y}\{V(h+d)\}=0 \quad [18]$$

where the coordinate system x is along the lines of longitude with east being in the positive direction, y is along the lines of latitude with south being in the positive direction, f is the Coriolis parameter, C_f is a nondimensional bottom frictional coefficient, and U and V are the depth-averaged velocities in the x and y directions, respectively. The velocities are averaged from the ocean bottom to the surface as

$$U=\frac{1}{(h+d)}\int_{-d}^{h}u\,dz=\frac{Q_x}{(h+d)}$$

$$V=\frac{1}{(h+d)}\int_{-d}^{h}v\,dz=\frac{Q_y}{(h+d)} \quad [19]$$

where Q_x and Q_y are flow rates in the x and y directions defined as $U(d+h)$ and $V(d+h)$, respectively. A typical value of the nondimensional frictional coefficient C_f is 10^{-3} for coastal water and 10^{-2} for run-ups on land (Satake, 1995).

4.19.6.6 Dispersion

This section briefly examines the general gravity waves deviating from the small-amplitude, or linear shallow-water (long), wave approximations (Mei, 1989; Peregrine, 1972). The phase velocity of a general gravity wave was previously shown by eqn [13]. If the hyperbolic tangent function is expanded into its Taylor series and the second term is included, eqn [13] becomes

$$c=\sqrt{gd}\left\{1-\frac{1}{3}(kd)^2\right\}^{1/2}\approx\sqrt{gd}\left\{1-\frac{1}{6}(kd)^2\right\} \quad [20]$$

The corresponding equation of motion is

$$\frac{\partial u}{\partial t}=-g\frac{\partial h}{\partial x}+\frac{1}{3}d^2\frac{\partial^3 u}{\partial x^2\partial t} \quad [21]$$

and the equation of continuity is

$$\frac{\partial h}{\partial t}=-d\frac{\partial u}{\partial x} \quad [22]$$

These equations, including the frequency dispersion, are known as the linear Boussinesq equation.

If the small-amplitude ($h\ll d$) assumption is relaxed, the nonlinear, nondispersive equations of motion and continuity are given as

$$\frac{\partial u}{\partial t}+u\frac{\partial u}{\partial x}=-g\frac{\partial h}{\partial x} \quad [23]$$

$$\frac{\partial h}{\partial t}=-\frac{\partial}{\partial x}\{u(h+d)\} \quad [24]$$

These equations are for the finite-amplitude shallow-water waves and account for the amplitude dispersion. For the linear case (ignoring the advection term), the phase velocity is given by $c=\sqrt{g(d+h)}$. This means that the larger the amplitude, the faster the wave speed. Consequently, the crest of a wave will catch up to the troughs in front of them, and the forward-facing portion of the wave will continue to steepen and will eventually break.

4.19.6.7 Numerical Computations

The tsunami propagation depends on the water depth, which is not uniform, but the variation is relatively well known compared to the velocity distribution of seismic waves. Thus, numerical computations on actual bathymetry have been popularly made. The historical review of tsunami numerical simulations was presented by Shuto (1991).

Let us consider a simple one-dimensional propagation (on constant depth) of linear long waves. Denoting the average velocity (x component only) as U, the equations of motion and continuity are given as

$$\frac{\partial U}{\partial t}=-g\frac{\partial h}{\partial x} \quad [25]$$

$$\frac{\partial h}{\partial t}=-d\frac{\partial U}{\partial x} \quad [26]$$

Using the staggered (leap-frog) grid system for U and h, and denoting time $t=l\,\Delta t$ as superscripts and space $x=m\,\Delta x$ as subscripts where l and m are integers, eqns [25] and [26] can be rewritten as

$$\frac{\left[U_m^{l+1/2}-U_m^{l-1/2}\right]}{\Delta t}=-g\frac{\left[h_{m+1/2}^{l}-h_{m-1/2}^{l}\right]}{\Delta x} \quad [27]$$

$$\frac{\left[h_{m+1/2}^{l+1}-h_{m+1/2}^{l}\right]}{\Delta t}=-d\frac{\left[U_{m+1}^{l+1/2}-U_m^{l+1/2}\right]}{\Delta x} \quad [28]$$

When using these equations, error will increase as time increases unless the following stability condition is met:

$$\sqrt{gd}\left(\frac{\Delta t}{\Delta x}\right)\le 1 \text{ or } \Delta t\le\frac{\Delta x}{\sqrt{gd}} \quad [29]$$

which is known as the Courant–Friedrichs–Lewy (CFL) condition. The physical meaning of the stability condition is that the time step Δt must be equal to or smaller than the time required for the disturbance to travel the spatial grid size Δx. If Δt is larger, the values other than $(l, m-1)$ $(l-1, m)$ and $(l, m+1)$ will influence the value at $(l+1, m)$. For the two-dimensional case, the CFL condition becomes

$$\Delta t\le\frac{\Delta x}{\sqrt{2gd}} \quad [30]$$

There are two types of boundaries involved in the tsunami computation. The first is a land–ocean boundary. The simplest assumption is a total reflection of energy on the coast. For the computation of tsunami inundation on land, however, a moving land–ocean boundary is needed; the boundary moves depending on the relative heights of the water level and land elevation. Several kinds of boundary conditions are described in Imamura (2009). The other kind of boundary is the open ocean boundary along the outer area of the computational region. For the radiation conditions, waves are assumed to leave the computational domain and maintain their slopes.

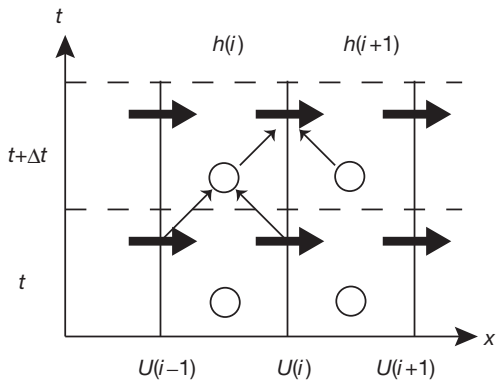


Figure 13 The staggered leap-frog system.

As an example of the numerical computation of tsunami waves, a numerical scheme is presented for a one-dimensional finite-difference computation. Since there is no overlap between the two (height and velocity) grids (Figure 13) in the linear case, in the program, the spatial grid points m and $m+\frac{1}{2}$ can be treated by the same index (e.g., i) and time $l-\frac{1}{2}$ and l can be treated as identical times (e.g., it). Writing U and h using the two-dimensional array as $U(i, \text{it})$ and $h(i, \text{it})$, eqns [27] and [28] become

$$U(i, \text{it}+1) = U(i, \text{it}) - g\frac{\Delta t}{\Delta x}[h(i, \text{it}) - h(i-1, \text{it})] \quad [31]$$

$$h(i, \text{it}+1) = h(i, \text{it}) - d\frac{\Delta t}{\Delta x}[U(i+1, \text{it}+1) - U(i, \text{it}+1)] \quad [32]$$

Thus, the algorithm becomes as follows:

```
set up parameters
read bathymetry data
read initial condition h(i,1)
while (it<itend)
 compute U(i,it+1) from U(i,it) and h(i,it)
 set U(i,it+1)=0 for land boundary
 compute h(i,it+1) from U(i,it+1) and h(i,it)
 adjust h(i,it+1) for open boundary
 exchange variables
  h(i,it+1)->h(i,it)
  U(i,it+1)->U(i,it)
 it+= 1
end
```

4.19.7 Tsunami Warning Systems

The 2004 Indian Ocean tsunami caused the worst tsunami disaster in history. In addition to the gigantic size of the earthquake, the lack of preparation for the tsunamis on the coasts of the Indian Ocean was also responsible for the tragedy. Preparation for a tsunami includes the use of a tsunami early warning system based on seismic and sea-level monitoring, hazard assessments such as making hazard maps or probabilistic estimates, and educational awareness of the coastal residents. Such systems for the reduction of tsunami disasters rely on the scientific knowledge of earthquakes and tsunamis.

The Pacific-wide tsunami warning system was established after the 1946 Aleutian tsunami that affected the Hawaiian Islands (Section 4.19.3.1.5). After the 1960 Chilean tsunami (Section 4.19.3.1.3), the Pacific tsunami warning system became international, and the International Coordination Group for the Tsunami Warning System in the Pacific (ICG/ITSU) was formed under the auspices of the UNESCO Intergovernmental Oceanographic Commission. Following the 2004 Indian Ocean tsunami, the group was renamed to the ICG for the Pacific Tsunami Warning and Mitigation System (ICG/PTWS), and similar groups were formed for the other ocean basins: ICG/IOTWS for the Indian Ocean, ICG/NEAMTWS for the northeastern Atlantic and Mediterranean Sea, and ICG/CARIBE-EWS for the Caribbean Sea.

In coordination with ICG/IOTWS, three regional tsunami warning centers were recently established in the countries located near the Indian Ocean. They are the German-Indonesian Tsunami Early Warning System (GITEWS) in Badan Meteorologi, Klimatologi, dan Geofisika (Munch et al., 2011), Joint Australian Tsunami Warning Centre (JATWC) of the Bureau of Meteorology and Geoscience Australia (Allen and Greenslade, 2008), and the Indian National Tsunami Early Warning System of the Indian Centre for Ocean Information Services (INCOIS) (Kumar et al., 2012). These centers are staffed 24 h a day and 7 days a week to monitor seismic activity and the possibility of a tsunami occurrence. The warning systems rely on seismic and sea-level monitoring, a database of past tsunami events, and premade numerical simulations. These systems typically issue tsunami warning messages about 5 min after an earthquake.

Tsunami warning systems can be grouped into basin-wide (such as the Pacific or Indian Ocean), regional, and local systems. Tsunamis travel very fast as an ocean wave with a speed of 710 km h^{-1} (0.2 km s^{-1}) at a water depth of 4000 m, but much slower than seismic waves, which travel at a speed of 5–10 km s^{-1}. This velocity difference between seismic and tsunami waves makes it possible to issue a tsunami warning by detecting seismic waves before the actual arrival of a tsunami. This lead time is more than a few hours for far-field tsunamis, while it can be as short as a few minutes for local tsunamis. Thus, far-field (basin-wide and regional) tsunami warning systems and local tsunami warning systems are designed differently, as described in the following sections.

4.19.7.1 Far-Field Tsunami Warning Systems

For distant tsunamis in the Pacific Ocean, the Richard H. Hagemeyer Pacific Tsunami Warning Center (PTWC), located in Hawaii, is the operational center for the international tsunami warning system and issues warnings for regional and distant tsunamis to almost all the countries and states surrounding the Pacific Ocean (McCreery, 2005). The West Coast/Alaska Tsunami Warning Center, located in Alaska, is responsible for issuing tsunami warnings along the US and Canadian coasts (excluding Hawaii). At these centers, seismic waveform data from several tens of stations are continuously monitored, and once a large earthquake is detected, the location, depth, and magnitudes are assessed using a combination

of automatic and interactive procedures by watch standers. The recent deployment of advanced seismological parameters, such as M_m (Okal and Talandier, 1989), M_{wp} (Tsuboi et al., 1995), Θ (Newman and Okal, 1998), and the W phase (Kanamori and Rivera, 2008), makes it possible to quickly assess an earthquake's size and tsunami potential and to issue tsunami warnings in less than half an hour. To avoid false alarms, it is important to confirm that a tsunami was actually generated; hence, sea levels are monitored via coastal tide gauges and tsunameters (DART buoys). If a tsunami is detected, then the tsunami warning is updated. If a tsunami is not confirmed, then the warning is canceled.

On 26 December 2004, during the 2004 Indian Ocean tsunami, the PTWC issued the first information bulletin at 1:14 GMT, only 15 min after the earthquake. An earthquake was located off the west coast of North Sumatra, and the magnitude was estimated to be 8.0 (M_{wp}). The second bulletin was issued at 2:08 GMT, 69 min after the earthquake, but still before the tsunami arrived at the coasts of Thailand, Sri Lanka, and India. The earthquake size was upgraded to 8.5 (M_m), and the possibility of a local tsunami was included in the bulletin.

A tsunami simulation database is used for data assimilation. At the NOAA Pacific Marine Environmental Laboratory or JATWC, tsunami waveforms at the DART locations are computed for numerous tsunami sources around the Pacific Ocean. When the DART stations record tsunami waveforms, they are compared with precomputed tsunami waveforms to first estimate the tsunami sources. Then, the estimated sources are used to predict tsunami arrival times and their amplitude at more distant locations. The predicted tsunami waveforms from the real-time data assimilation show good agreement with the observed waveforms (Titov et al., 2005b). Furthermore, tsunami inundation modeling for coastal communities such as Hawaii (Tang et al., 2009) or the amount of tsunami energy (Tang et al., 2012) is also made prior to the arrival of actual tsunamis.

4.19.7.2 Local Tsunami Warning Systems

For local tsunamis, time is more critical. In order to promptly issue tsunami warning messages, most near-field tsunami warning systems, including that of the JMA, rely on seismic observations to estimate the location, depth, and size of an earthquake. At the JMA, the epicenter and magnitude M of earthquakes are first estimated within a few minutes of occurrence using a seismographic network that consists of several hundreds of stations, and the likelihood of a tsunami occurrence is judged from these parameters (Kamigaichi, 2009) (**Figure 14**). The JMA's tsunami warning system consists of three categories: tsunami warnings for major tsunamis (estimated tsunami height >3 m), tsunami warnings (1 m <tsunami height <3 m), and tsunami advisories (0.2 m <tsunami height <1 m).

The JMA introduced a numerical simulation technique in 1999 (Tatehata, 1997); the generation and propagation for 100 000 different tsunami cases were calculated, and the results have been stored in a database. When a large earthquake occurs, the most appropriate case for the actual location and magnitude is retrieved from the database, and the expected tsunami arrival times and heights for the 66 segments of the Japanese coast are announced. The tsunami warning messages are updated based on more comprehensive seismic observations, and the confirmation of a tsunami is based on sea-level measurements.

In 1950s, when the JMA first introduced the tsunami warning system, it took about 20 min to issue a tsunami warning. In 1983, when a large earthquake occurred in the Sea of Japan, the warning was issued 12 min after the earthquake, but the tsunami arrived 7 min after the earthquake, which resulted in 100 casualties. They improved this warning system, and in 1993, a tsunami warning was issued 5 min after the southwest Hokkaido earthquake. However, the tsunami arrived earlier, which resulted in 230 casualties (**Section 4.19.3.1.2**). They further improved the system and now issue a tsunami warning within 3–5 min after large earthquakes occur.

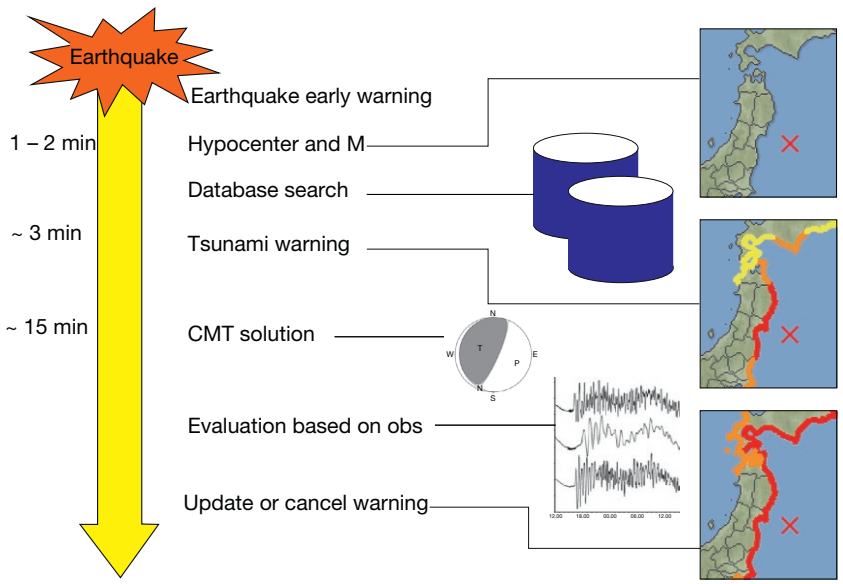


Figure 14 The system and timeline of the tsunami warning system of the Japan Meteorological Agency.

For the 2011 Tohoku earthquake, the JMA issued the first tsunami warning 3 min after the earthquake. The estimated tsunami heights were 6 m on the Miyagi coast and 3 m or less for the other coasts (Ozaki, 2011). These heights were significantly underestimated compared to the actual tsunami heights (Section 4.19.2.2) because the initial estimate of magnitude ($M=7.9$) was much smaller than the actual size ($M_w=9.0$). The earthquake size, the moment magnitude M_w, is usually reestimated using broadband seismograph data within 15 min. However, because of the gigantic size of the Tohoku earthquake, most domestic broadband seismograms went off-scale and the JMA was not able to estimate M_w using the Japanese data. It took almost 1 h for the JMA to tentatively estimate the moment magnitude of the earthquake ($M_w=8.8$) after using foreign seismographic observations.

After detecting the large offshore tsunami via GPS wave gauges (Figure 5), the JMA upgraded the tsunami warning messages to a higher level of estimated tsunami heights 28 min after the earthquake. Although the first update was before the actual arrival of the tsunami and updates were made several times, the updated tsunami warning messages did not reach all the coastal communities due to power failures and the fact that many people had already started to evacuate.

Based on lessons from the 2011 Tohoku earthquake, the JMA has made efforts to improve its tsunami warning system. Technical developments have been included to estimate the earthquake size within a few minutes by using a suite of information, with some being redundant, and by deploying and utilizing the offshore tsunami observation systems. In terms of dissemination, if the earthquake size is potentially underestimated, the JMA issues a warning based on a worst case scenario. The expected tsunami heights will then be announced by simple and qualitative expressions considering the uncertainties.

For local tsunamis, earthquake ground shaking may be the best natural warning. When a strong earthquake is felt near the coast, one should run to high ground or to the higher levels of a strong structure. Such behavior saved thousands of lives during the 2011 Tohoku tsunami. However, the weak shaking of a tsunami earthquake may be also followed by a large tsunami, as demonstrated by the 2010 Mentawai earthquake.

4.19.8 Tsunami Hazard Assessments

4.19.8.1 Tsunami Hazard Maps

Once coastal residents receive a tsunami warning message, they need to know what it means and where to evacuate. An effective tool for public use is a map showing the area to evacuate for future tsunamis. Possible inundation zones are estimated either from historical data of past tsunamis or by tsunami numerical simulation. Safe places such as tsunami shelters can also be shown on tsunami hazard maps. Tsunami hazard maps help coastal communities prepare for tsunami hazards.

Tsunami hazard maps can be constructed in either a deterministic way (deterministic tsunami hazard assessment, DTHA) or a probabilistic way (probabilistic tsunami hazard assessment, PTHA). In the DTHA, a maximum credible or most likely tsunami source is determined, and then, the tsunami propagation and coastal inundation are calculated. Uncertainty associated with the estimation of the source parameters must also be considered. The PTHA will be described in the next Section 4.19.8.2. In the United States, the NOAA National Tsunami Hazard Mitigation Program develops and updates tsunami hazard maps using the latest tsunami inundation simulation. On the basis of tsunami hazard maps, evacuation plans can be established by responsible officials, such as civil defense officials (González et al., 2005b). One of the lessons learned from the 2004 Indian Ocean tsunami is that both coastal residents and foreign tourists need to be informed about tsunami hazards. For the Hawaiian Islands, tsunami hazard maps are prepared for most of the occupied coasts and published in the local phone books that are available at every hotel room.

In Japan, tsunami hazard maps were printed and distributed to coastal residents before the 2011 Tohoku tsunami in most of the affected coastal communities. Tsunami inundation areas have been estimated from past tsunami data or computed using a scenario of an earthquake. On the Sanriku coast, most hazard maps show the inundation areas from past tsunamis, typically the 1896 Sanriku tsunami. Because the 1896 and 2011 tsunami heights and inundation areas were roughly similar for most of the Sanriku coast, the 2011 tsunami inundation areas were also similar to those shown in the hazard maps. On the contrary, in the Sendai plain, tsunami hazard maps and other countermeasures were prepared for a Miyagi-oki earthquake of $M\sim8$, which was estimated to occur with 99% probability in the next 30 years. The predicted inundation area for a typical $M\sim8$ earthquake was, for the most part, within 1 km from the coast and much smaller than the actual tsunami inundation area of the 2011 $M\sim9$ earthquake (Figure 15). The distribution of the 869 Jogan tsunami deposits, however, was similar to the inundation area of the 2011 Tohoku tsunami.

4.19.8.2 Probabilistic Tsunami Hazard Assessment

The PTHA has been formulated in an analogous way to that of the probabilistic seismic hazard assessment (Cornell, 1968; Chapter 4.23). Results of the PTHA are typically displayed as hazard curves that show the annual frequency of exceedance of run-up height values. The hazards from a large number of possible sources are aggregated together to develop a tsunami hazard curve. In addition to including many different types of tsunami generators, multiple sources of uncertainty related to the source parameters and numerical models of tsunamis are considered in the PTHA. Two types of uncertainty, aleatory and epistemic, are distinguished. Aleatory uncertainty, or random variability, relates to the natural or stochastic uncertainty inherent in a physical system. Aleatory uncertainty is typically represented by a probability distribution from repeated observations or experiments. Epistemic uncertainty is due to incomplete knowledge and data and can be reduced by the collection of new data. Uncertainties in various model parameters and alternative techniques and characterizations are treated as logic trees (Annaka et al., 2007). A single hazard curve is obtained by integration over the aleatory uncertainties and all possible sources. A large number of hazard curves are obtained for different branches of a logic tree representing epistemic uncertainty.

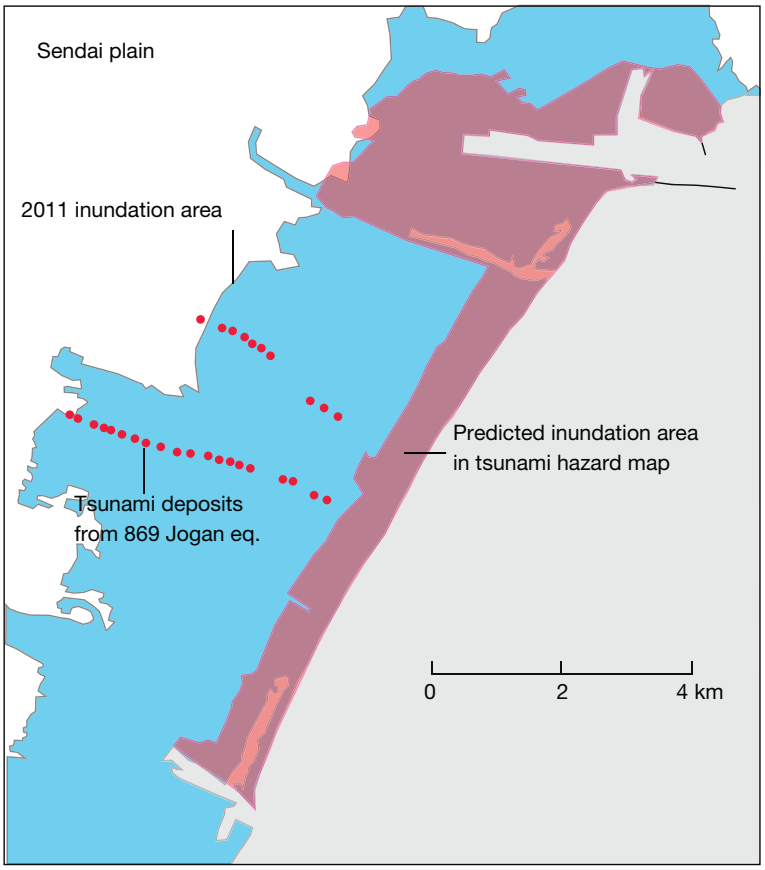


Figure 15 Tsunami hazard map for the Sendai plain (based on Central Disaster Management Council, 2011). The expected tsunami inundation zone assuming a Miyagi-oki earthquake ($M = 8.0$) is shown by brown color. The actual tsunami inundation area of the 2011 Tohoku earthquake is shown by the light blue color. The distribution of the 869 Jogan tsunami deposits (Sawai et al., 2012) is shown by red circles.

The tsunami hazard probabilities can be estimated from historical data. If we know that a particular coast has experienced N tsunamis within a period of time T, assuming a Poisson process, the probability P of having at least one tsunami for a future time period of τ is given as

$$P = 1 - \exp(-N_\tau/T) \qquad [33]$$

For example, if two tsunamis were recorded in the last 100 years at a particular location, the probability of having another tsunami in the next 10 years is 18%. If the tsunami heights from past events are known, the probabilities for each range of tsunami heights can be similarly computed (Geist and Parsons, 2006; Parsons and Geist, 2008) (**Figure 16**).

Great earthquakes that have occurred at plate boundaries may recur more regularly than the Poisson process. For the Poisson process, the time interval between earthquakes would follow a power law without a particular peak. The time interval between large interpolate earthquakes can be modeled using Weibull or a log-normal probability distribution (Nishenko and Buland, 1987; Rikitake, 1976; Utsu, 1984). The Brownian passage-time distribution has been recently used since it models steady tectonic loading plus the Brownian perturbation (Matthews et al., 2002). Then, the probability of the next earthquake in a specific time window can be computed if the average recurrence period and date of the most recent event, as well as the aperiodicity parameter or variation coefficient, are known.

When the fault parameters of a characteristic tsunamigenic earthquake are known, tsunami heights can be estimated by numerical computations, and the probabilistic tsunami heights can be calculated. Such a probabilistic estimate of tsunami heights was made in the 1980s for the Japanese coasts (Rikitake and Aida, 1988). For the Pacific coast of Hokkaido, where the 2003 Tokachi-oki earthquake actually generated the tsunami, the probability of a 2 m tsunami between 2000 and 2010 was estimated as 7–15% and that of a 5 m tsunami was 7%. The actual tsunami heights in 2003 were up to 4 m (Tanioka et al., 2004). For the Sanriku and Fukushima coasts, where the tsunami heights from the 2011 Tohoku tsunami exceeded 10 m, the probability of such a large tsunami was estimated as zero for the 10-year period, while the probability for smaller (1 m or larger) tsunamis was estimated as 15–39%.

4.19.9 Summary

As demonstrated in this chapter, tsunami science has interdisciplinary natures. Tsunami sources include giant earthquakes, such as the 2004 Sumatra–Andaman earthquake, the 2011 Tohoku earthquake, and unusual tsunami earthquakes. More

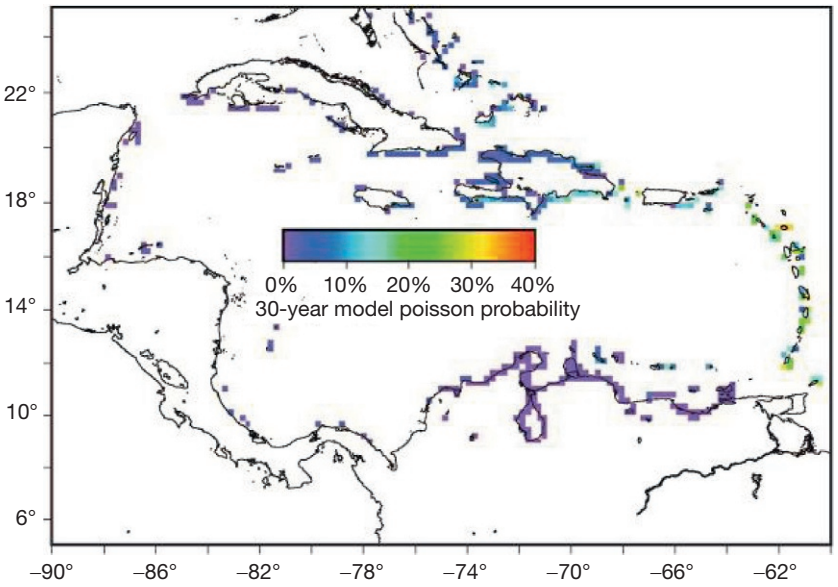


Figure 16 An example of the probabilistic tsunami hazard assessment. The 30-year probabilities of a tsunami run-up of 0.5 m or higher are shown in different colors. Reproduced from Parsons T and Geist E (2008) Tsunami probability in the Caribbean Region. *Pure and Applied Geophysics* 165: 2089–2116, with permission.

seismological studies are needed to better understand the generation process and mechanism of such infrequent earthquakes. Tsunami generation mechanisms from landslides or volcanic eruptions are not very well understood. While tsunami propagation in a deep ocean has been popularly treated by shallow-water theory and numerical computations, the observations and monitoring of offshore tsunami gauges or satellite altimeters have recently become possible. Such offshore observations in the future would provide real-time data for tsunami warnings and scientific data to examine the propagation characteristics of a tsunami, such as the dispersion effects or the coupling between ocean and the elastic Earth. The nearshore and onshore behaviors of a tsunami need to be studied more using theory, field data, and physical experiments. Geologic studies of past tsunamis are also important, as tsunamis are infrequent phenomena.

The 2004 Indian Ocean tsunami demonstrated that the tsunami effect can be global. For the reduction of future tsunami damage, it is essential to educate the global citizens about tsunamis. After the 1960 Chilean tsunami, the scientific knowledge of tsunamis has significantly improved, and we now can detect the occurrence of an earthquake and assess the possibility of a tsunami occurrence in real time. The 2011 Tohoku tsunami indicated that there are still more technical challenges to improve the tsunami warning system. Such scientific and technological developments can be and should be used to save human lives throughout the world.

References

Abe K (1973) Tsunami and mechanism of great earthquakes. *Physics of the Earth and Planetary Interiors* 7: 143–153.

Abe K (1979) Size of great earthquakes of 1873–1974 inferred from tsunami data. *Journal of Geophysical Research: Solid Earth* 84: 1561–1568.

Abe K (1981) Physical size of tsunamigenic earthquakes of the northwestern Pacific. *Physics of the Earth and Planetary Interiors* 27: 194–205.

Abe K (1989) Quantification of tsunamigenic earthquakes by the Mt scale. *Tectonophysics* 166: 27–34.

Abe K (2000) Predominance of long periods in large Pacific tsunamis. *Science of Tsunami Hazards* 18: 15–34.

Aida I (1978) Reliability of a tsunami source model derived from fault parameters. *Journal of Physics of the Earth* 26: 57–73.

Allen SCR and Greenslade DJM (2008) Developing tsunami warnings from numerical model output. *Natural Hazards* 46: 35–52.

Ammon CJ, Ji C, Thio HK, et al. (2005) Rupture process of the 2004 Sumatra–Andaman earthquake. *Science* 308: 1133–1139.

Annaka T, Satake K, Sakakiyama T, et al. (2007) Logic-tree approach for probabilistic tsunami hazard analysis and its applications to the Japanese coasts. *Pure and Applied Geophysics* 164: 577–592.

Atwater BF, Musumi-Rokkaku S, Satake K, et al. (2005) *The orphan Tsunami of 1700. U. S. G. S. Professional Paper 1707*. 133 pp.

Baptista MA, Heitor S, Miranda JM, et al. (1998) The 1755 Lisbon tsunami: Evaluation of the tsunami parameters. *Journal of Geodynamics* 25: 143–157.

Bardet JP, Synolakis CE, Davies HL, et al. (2003) Landslide tsunamis: Recent findings and research directions. *Pure and Applied Geophysics* 160: 1793–1809.

Bernard EN and Robinson AR (eds.) (2009) Tsunamis. In: *The Sea*, vol. 15. Cambridge: Harvard University Press.

Berninghausen WH (1966) Tsunamis and seismic seiches reported from regions adjacent to the Indian Oceans. *Bulletin of the Seismological Society of America* 56: 69–74.

Bilek SL and Lay T (2002) Tsunami earthquakes possibly widespread manifestations of frictional conditional stability. *Geophysical Research Letters* 29. http://dx.doi.org/10.1029/2002GL015215.

Bilham R (2005) A flying start, then a slow slip. *Science* 308: 1126–1127.

Bondevik S, Lovholt F, Harbitz C, et al. (2005) The Storegga slide tsunami – Comparing field observations with numerical simulations. *Marine and Petroleum Geology* 22: 195–208.

Borrero JC (2005) Field survey of Northern Sumatra and Banda Aceh, Indonesia after the tsunami and earthquake of 26 December 2004. *Seismological Research Letters* 76: 312–320.

Bourgeois J (2009) Geological effects and records of tsunamis. In: Bernard EN and Robinson AL (eds.) *Tsunamis*. *The Sea*, vol. 15, pp. 55–91. Cambridge, MA: Harvard University Press.

Bourgeois J, Hansen TA, Wiberg PL, et al. (1988) A tsunami deposit at the Cretaceous–Tertiary boundary in Texas. *Science* 241: 567–570.

Bryant EA, Young RW, and Price DM (1996) Tsunami as a major control on coastal evolution, southeastern Australia. *Journal of Coastal Research* 12: 831–840.

Bugge T, Belderson RH, and Kenyon NH (1988) The Storegga slide. *Philosophical Transactions of the Royal Society A* 325: 357–388.

Central Disaster Management Council (2011) Report of the Committee for Technical Investigation on countermeasures for earthquakes and tsunamis based on the lessons learned from the "2011 off the Pacific coast of Tohoku Earthquake", Tokyo. http://www.bousai.go.jp/jishin/chubou/higashinihon/Report.pdf.

Cisternas M, Atwater BF, Torrejon F, et al. (2005) Predecessors of the giant 1960 Chile earthquake. *Nature* 437: 404–407.

Comer RP (1980) Tsunami height and earthquake magnitude: theoretical basis of an empirical relation. *Geophysical Research Letters* 7: 445–448.

Comer RP (1984a) Tsunami generation: A comparison of traditional and normal mode approaches. *Geophysical Journal of the Royal Astronomical Society* 77: 29–41.

Comer RP (1984b) The tsunami mode of a flat earth and its excitation by earthquake sources. *Geophysical Journal of the Royal Astronomical Society* 77: 1–27.

Cornell CA (1968) Engineering seismic risk analysis. *Bulletin of the Seismological Society of America* 58: 1583–1606.

Cummins P (2004) Small threat, but warning sounded for tsunami research. *AusGeo News* 75: 4–7.

Dawson AG, Long D, and Smith DE (1988) The Storegga slides—Evidence from eastern Scotland for a possible tsunami. *Marine Geology* 82: 271–276.

Dawson AG and Shi SZ (2000) Tsunami deposits. *Pure and Applied Geophysics* 157: 875–897.

Dunbar P and McCullough H (2012) Global tsunami deposits database. *Natural Hazards* 63: 267–278.

Dunbar P, Stroker K, Brocko V, et al. (2008) Long-term tsunami data archive supports tsunami forecast, warning, research, and mitigation. *Pure and Applied Geophysics* 165: 2275–2291.

Fine IV, Rabinovich AB, Bornhold BD, et al. (2005) The Grand Banks landslide-generated tsunami of November 18, 1929: Preliminary analysis and numerical modeling. *Marine Geology* 215: 45–57.

Fine IV, Rabinovich AB, and Thomson RE (2005) The dual source region for the 2004 Sumatra tsunami. *Geophysical Research Letters* 32. http://dx.doi.org/10.1029/2005GL023521.

Fine IV, Rabinovich AB, Thomson RE, et al. (2003) Numerical modeling of tsunami generation by submarine and subaerial landslides. In: Yalçıner AC, Pelinovsky E, Okal E, and Synolakis C (eds.) *Submarine Landslides and Tsunamis*, pp. 69–88. The Netherlands: Kluwer.

Fritz HM, Mohammed F, and Yoo J (2009) Lituya Bay landslide impact generated mega-tsunami 50th anniversary. *Pure and Applied Geophysics* 166: 153–175.

Fritz H, Petroff C, Catalán P, et al. (2011) Field survey of the 27 February 2010 Chile tsunami. *Pure and Applied Geophysics* 168: 1989–2010.

Fujii Y and Satake K (2006) Source of the July 2006 West Java tsunami estimated from tide gauge records. *Geophysical Research Letters* 33. http://dx.doi.org/10.1029/2006GL028049.

Fujii Y and Satake K (2007) Tsunami source model of the 2004 Sumatra–Andaman earthquake inferred from tide gauge and satellite data. *Bulletin of the Seismological Society of America* 97: S192–S207.

Fujii Y and Satake K (2008) Tsunami waveform inversion of the 2007 Bengkulu, southern Sumatra earthquake. *Earth, Planets and Space* 60: 993–998.

Fujii Y and Satake K (2013) Slip distribution and seismic moment of the 2010 and 1960 Chilean earthquakes inferred from tsunami waveforms and coastal geodetic data. *Pure and Applied Geophysics* 170: 1493–1509.

Fujii Y, Satake K, Sakai S, et al. (2011) Tsunami source of the 2011 off the Pacific coast of Tohoku earthquake. *Earth, Planets and Space* 63: 815–820.

Fujino S, Naruse H, Matsumoto D, et al. (2009) Stratigraphic evidence for pre-2004 tsunamis in southwestern Thailand. *Marine Geology* 262: 25–28.

Fujiwara T, Kodaira S, No T, et al. (2011) The 2011 Tohoku-oki earthquake: displacement reaching the trench axis. *Science* 334: 1240.

Fukao Y (1979) Tsunami earthquakes and subduction processes near deep-sea trenches. *Journal of Geophysical Research: Solid Earth* 84: 2303–2314.

Geist E (1998) Local tsunamis and earthquake source parameters. *Advances in Geophysics* 39: 117–209.

Geist E (2000) Origin of the 17 July 1998 Papua New Guinea tsunami: Earthquake or landslide? *Seismological Research Letters* 71: 344–351.

Geist E and Parsons T (2006) Probabilistic analysis of tsunami hazards. *Natural Hazards* 37: 277–314.

Gelfenbaum G and Jaffe B (2003) Erosion and sedimentation from the 17 July, 1998 Papua New Guinea tsunami. *Pure and Applied Geophysics* 160: 1969–1999.

Goff J, Chagué-Goff C, Nichol S, et al. (2012) Progress in palaeotsunami research. *Sedimentary Geology* 243–244: 70–88.

González FI, Bernard EN, Meinig C, et al. (2005) The NTHMP tsunameter network. *Natural Hazards* 35: 25–39.

González FI, Titov VV, Mofjeld HO, et al. (2005) Progress in NTHMP hazard assessment. U.S. National Tsunami Hazard Mitigation Program. *Natural Hazards* 35: 89–110.

Goto K, Chagué-Goff C, Fujino S, et al. (2011) New insights of tsunami hazard from the 2011 Tohoku-oki event. *Marine Geology* 290: 46–50.

Goto K, Kawana T, and Imamura F (2010) Historical and geological evidence of boulders deposited by tsunamis, southern Ryukyu Islands, Japan. *Earth-Science Reviews* 102: 77–99.

Goto K, Tada R, Tajika E, et al. (2008) Lateral lithological and compositional variations of the Cretaceous/Tertiary deep-sea tsunami deposits in northwestern Cuba. *Cretaceous Research* 29: 217–236.

Gower J (2005) Jason-1 detects the 26 December 2004 tsunami. *Eos, Transactions American Geophysical Union* 86: 37–38.

Grilli ST and Watts P (2005) Tsunami generation by submarine mass failure I: Modeling, experimental validation, and sensitivity analyses. *Journal of Waterway, Port, Coastal, and Ocean Engineering* 131: 283–297.

Gusiakov VK (2005) Tsunami generation potential of different tsunamigenic regions in the Pacific. *Marine Geology* 215: 3–9.

Hanks T and Kanamori H (1979) A moment magnitude scale. *Journal of Geophysical Research: Solid Earth* 84: 2348–2350.

Hanson JF, Reasoner C, and Bowman JR (2007) High frequency tsunami signals of the great Indonesian earthquakes of 26 December and 28 March 2005. *Bulletin of the Seismological Society of America* 97: S232–S248.

Harbitz CB (1992) Model simulations of tsunamis generated by the Storegga slides. *Marine Geology* 105: 1–21.

Hatori T (1979) Relation between tsunami magnitude and wave energy. *Bulletin of the Earthquake Research Institute, University of Tokyo* 54: 531–541.

Hatori T (1984) On the damage to houses due to tsunamis. *Bulletin of the Earthquake Research Institute, University of Tokyo* 59: 433–439.

Heezen BC and Ewing M (1952) Turbidity currents and submarine slumps, and the 1929 Grand Banks earthquake. *American Journal of Science* 250: 849–873.

Heidarzadeh M and Satake K (2013) Waveform and spectral analyses of the 2011 Japan tsunami records on tide gauge and DART stations across the Pacific Ocean. *Pure and Applied Geophysics* 170: 1275–1293.

Heinrich P, Piatanesi A, and Hebert H (2001) Numerical modelling of tsunami generation and propagation from submarine slumps: the 1998 Papua New Guinea event. *Geophysical Journal International* 145: 97–111.

Hirata K, Aoyagi M, Mikada H, et al. (2002) Real-time geophysical measurements on the deep seafloor using submarine cable in the southern Kurile subduction zone. *IEEE Journal of Oceanic Engineering* 27: 170–181.

Hirata K, Satake K, Tanioka Y, et al. (2006) The 2004 Indian Ocean tsunami: Tsunami source model from satellite altimetry. *Earth, Planets and Space* 58: 195–201.

Iida K (1984) *Catalog of Tsunamis in Japan and its Neighboring Countries,* 52 pp. Toyota: Aichi Institute of Technology.

Iida K, Cox DC, and Pararas-Carayannis G (1967) *Preliminary Catalog of Tsunamis Occurring in the Pacific Ocean,* 274 pp. Honolulu: University of Hawaii.

Imamura F (2009) Tsunami modeling: Calculating inundation and hazard maps. In: Bernard EN and Robinson AR (eds.) *Tsunamis*. *The Sea*, vol. 15, pp. 321–332. Cambridge: Harvard University Press.

Imamura F and Hashi K (2003) Re-examination of the source mechanism of the 1998 Papua New Guinea earthquake and tsunami. *Pure and Applied Geophysics* 160: 2071–2086.

Intergovernmental Oceanographic Commission (2006) *Manual on Sea Level Measurement and Interpretation Volume IV: An Update to 2006*, 78 pp. Paris: UNESCO.

Intergovernmental Oceanographic Commission (2013) *Post-Tsunami Survey Field Guide (Second Edition)*, 89 pp. Paris: UNESCO.

International Federation of Red Cross and Red Crescent Societies (2005) *World Disasters Report*, 251 pp. Geneva: Kumarian Press.

Ishii M, Shearer PM, Houston H, et al. (2005) Extent, duration and speed of the 2004 Sumatra–Andaman earthquake imaged by the Hi-Net array. *Nature* 435: 933–936.

Ito Y, Tsuji T, Osada Y, et al. (2011) Frontal wedge deformation near the source region of the 2011 Tohoku-oki earthquake. *Geophysical Research Letters* 38. http://dx.doi.org/10.1029/2011GL048355.

Iwasaki SI, Eguchi T, Fujinawa Y, et al. (1997) Precise tsunami observation system in the deep ocean by an ocean bottom cable network for the prediction of earthquakes and tsunamis. In: Hebenstreit G (ed.) *Perspectives on Tsunami Hazard Reduction*, pp. 47–66. Dordrecht: Kluwer Academic Publishers.

Jaffe BE, Borrero JC, Prasetya GS, et al. (2006) The December 2004 Indian Ocean tsunami in northwest Sumatra and offshore islands. *Earthquake Spectra* 22: S105–S135.

Jankaew K, Atwater BF, Sawai Y, et al. (2008) Medieval forewarning of the 2004 Indian Ocean tsunami in Thailand. *Nature* 455: 1228–1231.
Jiang L and Leblond PH (1994) 3-Dimensional modeling of tsunami generation due to a submarine mudslide. *Journal of Physical Oceanography* 24: 559–572.
Johnson JM (1999) Heterogeneous coupling along Alaska-Aleutians as inferred from tsunami, seismic and geodetic inversions. *Advances in Geophysics* 39: 1–116.
Johnson JM and Satake K (1997) Estimation of seismic moment and slip distribution of the April 1, 1946, Aleutian tsunami earthquake. *Journal of Geophysical Research: Solid Earth* 102: 11765–11774.
Kajiura K (1963) The leading wave of a tsunami. *Bulletin of the Earthquake Research Institute, University of Tokyo* 41: 535–571.
Kajiura K (1970) Tsunami source, energy and the directivity of wave radiation. *Bulletin of the Earthquake Research Institute, University of Tokyo* 48: 835–869.
Kamigaichi O (2009) Tsunami forecasting and warning. In: Meyers RA (ed.) *Encyclopedia of Complexity and Systems Science*, pp. 9592–9613. New York: Springer.
Kanamori H (1971) Seismological evidence for a lithospheric normal faulting—The Sanriku earthquake of 1933. *Physics of the Earth and Planetary Interiors* 4: 289–300.
Kanamori H (1972) Mechanism of tsunami earthquakes. *Physics of the Earth and Planetary Interiors* 6: 246–259.
Kanamori H and Kikuchi M (1993) The 1992 Nicaragua earthquake: A slow tsunami earthquake associated with subducted sediments. *Nature* 361: 714–716.
Kanamori H and Rivera L (2008) Source inversion of W phase: Speeding up seismic tsunami warning. *Geophysical Journal International* 175: 222–238.
Kato T, Terada Y, Kinoshita M, et al. (2000) Real-time observation of tsunami by RTK-GPS. *Earth, Planets and Space* 52: 841–845.
Kato T, Terada Y, Nishimura H, et al. (2011) Tsunami records due to the 2010 Chile earthquake observed by GPS buoys established along the Pacific coast of Japan. *Earth, Planets and Space* 63: e5–e8.
Kawamata K, Takaoka K, Ban K, et al. (2005) Model of tsunami generation by collapse of volcanic eruption: The 1741 Oshima–Oshima tsunami. In: Satake K (ed.) *Tsunamis: Case Studies and Recent Developments*, pp. 79–96. Dordrecht: Springer.
Kawata Y, Benson BC, Borrero JC, et al. (1999) Tsunami in Papua New Guinea was as intense as first thought. *Eos, Transactions American Geophysical Union* 80(101): 105.
Kayanne H, Ikeda Y, Echigo T, et al. (2007) Coseismic and postseismic creep in the Andaman Islands associated with the 2004 Sumatra–Andaman earthquake. *Geophysical Research Letters* 34: http://dx.doi.org/10.129/2006GL028200.
Keating BH, Waythomas CF, and Dawson AG (2000) Special issue: Landslides and tsunamis—Introduction. *Pure and Applied Geophysics* 157: 871–873.
Keller G (2011) The Cretaceous–Tertiary mass extinction: Theories and controversies. *Special Publication: Society for Sedimentary Geology* 100: 7–22.
Kido M, Osada Y, Fujimoto H, et al. (2011) Trench-normal variation in observed seafloor displacements associated with the 2011 Tohoku-oki earthquake. *Geophysical Research Letters* 38. http://dx.doi.org/10.1029/2011GL050057.
Koshimura S, Oie T, Yanagisawa H, et al. (2009) Developing fragility functions for tsunami damage estimation using numerical model and post-tsunami data from Banda Aceh, Indonesia. *Coastal Engineering Journal* 51: 243–273.
Kulikov EA, Rabinovich AB, Thomson RE, et al. (1996) The landslide tsunami of November 3, 1994, Skagway Harbor, Alaska. *Journal of Geophysical Research: Oceans* 101: 6609–6615.
Kumar TS, Nayak S, Kumar P, et al. (2012) Performance of the tsunami forecast system for the Indian Ocean. *Current Science* 102: 110–114.
Lander JF (1996) *Tsunamis Affecting Alaska: 1737–1996*, 195 pp. Boulder, CO: NOAA National Geophysical Data Center,
Lander JF and Lockridge PA (1989) *United States Tsunamis*, 265 pp. Boulder, CO: NOAA National Geophysical Data Center.
Lander JF, Lockridge PA, and Kozuch MJ (1993) *Tsunamis Affecting the West Coast of the United States 1806–1992*, 242 pp. Boulder, CO: NOAA National Geophysical Data Center.
Latter JH (1981) Tsunamis of volcanic origin: Summary of causes, with particular reference to Krakatoa, 1883. *Bulletin of Volcanology* 44: 467–490.
Lavigne F, Paris R, Grancher D, et al. (2009) Reconstruction of tsunami inland propagation on December 26, 2004 in Banda Aceh, Indonesia, through field investigations. *Pure and Applied Geophysics* 166: 259–281.
Lay T, Kanamori H, Ammon CJ, et al. (2005) The great Sumatra–Andaman earthquake of 26 December 2004. *Science* 308: 1127–1133.
Liu PLF, Lynett P, Fernando H, et al. (2005) Observations by the International Tsunami Survey Team in Sri Lanka. *Science* 308: 1595.
Maeno F and Imamura F (2011) Tsunami generation by a rapid entrance of pyroclastic flow into the sea during the 1883 Krakatau eruption, Indonesia. *Journal of Geophysical Research: Solid Earth* 116. http://dx.doi.org/10.1029/2011JB008253.
Malik JN and Murty CVR (2005) Landscape changes in Andaman and Nicobar Islands (India) due to Mw 9.3 tsunamigenic Sumatra earthquake of 26 December 2004. *Current Science* 88: 1384–1386.
Malik JN, Shishikura M, Echigo T, et al. (2011) Geologic evidence for two pre-2004 earthquakes during recent centuries near Port Blair, South Andaman Island, India. *Geology* 39: 559–562.
Mansinha L and Smylie DE (1971) The displacement fields of inclined faults. *Bulletin of the Seismological Society of America* 61: 1433–1440.
Mas E, Koshimura S, Suppasri A, et al. (2012) Developing Tsunami fragility curves using remote sensing and survey data of the 2010 Chilean Tsunami in Dichato. *Natural Hazards and Earth System Sciences* 242: 2689–2697.
Matsuyama M, Walsh JP, and Yeh H (1999) The effect of bathymetry on tsunami characteristics at Sisano Lagoon, Papua New Guinea. *Geophysical Research Letters* 26: 3513–3516.
Matthews MV, Ellsworth WL, and Reasenberg PA (2002) A Brownian model for recurrent earthquakes. *Bulletin of the Seismological Society of America* 92: 2233–2250.
McCreery CS (2005) Impact of the National Tsunami Hazard Mitigation Program on operations of the Richard H. Hagemeyer Pacific Tsunami Warning Center. *Natural Hazards* 35: 73–88.
Mei CC (1989) *The Applied Dynamics of Ocean Surface Waves*, 740 pp. Singapore: World Scientific.
Meltzner AJ, Sieh K, Abrams M, et al. (2006) Uplift and subsidence associated with the great Aceh–Andaman earthquake of 2004. *Journal of Geophysical Research: Solid Earth* 111. http://dx.doi.org/10.1029/2005JB003891.
Merrifield MA, Firing YL, Aarup T, et al. (2005) Tide gauge observations of the Indian Ocean tsunami, December 26, 2004. *Geophysical Research Letters* 32. http://dx.doi.org/10.1029/2005GL022610.
Miller DJ (1960) *Giant waves in Lituya Bay, Alaska. U. S. G. S. Professional Paper* 354-C, pp. 51–86.
Minoura K, Imamura F, Sugawara D, et al. (2001) The 869 Jogan tsunami deposit and recurrence interval of large-scale tsunami on the Pacific coast of northeast Japan. *Journal of Natural Disaster Science* 23: 83–88.
Minoura K and Nakaya S (1991) Traces of tsunami preserved in inter-tidal lacustrine and marsh deposits—Some examples from northeast Japan. *The Journal of Geology* 99: 265–287.
Miyabe N (1934) An investigation of the Sanriku tsunami based on mareogram data. *Bulletin of the Earthquake Research Institute, University of Tokyo* (supplement 1): 112–126.
Monecke K, Finger W, Klarer D, et al. (2008) A 1000-year sediment record of tsunami recurrence in northern Sumatra. *Nature* 455: 1232–1234.
Moore A, Nishimura Y, Gelfenbaum G, et al. (2006) Sedimentary deposits of the 26 December 2004 tsunami on the northwest. *Earth, Planets and Space* 58: 253–258.
Mori N, Takahashi T, Yasuda T, et al. (2011) Survey of 2011 Tohoku earthquake tsunami inundation and run-up. *Geophysical Research Letters* 38. http://dx.doi.org/10.1029/2011GL049210.
Morton RA, Gelfenbaum G, Buckley ML, et al. (2011) Geological effects and implications of the 2010 tsunami along the central coast of Chile. *Sedimentary Geology* 242: 34–51.
Morton RA, Gelfenbaum G, and Jaffe BE (2007) Physical criteria for distinguishing sandy tsunami and storm deposits using modern examples. *Sedimentary Geology* 200: 184–207.
Munch U, Rudloff A, and Lauterjung J (2011) Postface "The GITEWS Project—Results, summary and outlook" *Natural Hazards and Earth System Sciences* 11: 765–769.
Murty TS and Rafiq M (1991) A tentative list of tsunamis in the marginal seas of the North Indian Ocean. *Natural Hazards* 4: 81–83.
Namegaya Y and Satake K (2013) Reexamination of the AD 869 Jogan earthquake size from tsunami deposit distribution, simulated flow depth, and velocity. *Geophysical Research Letters* 41. http://dx.doi.org/10.1002/2013GL058678.
Nagarajan B, Suresh I, Sundar D, et al. (2006) The great tsunami of 26 December 2004: A description based on tide-gauge data from the Indian subcontinent and surrounding areas. *Earth, Planets and Space* 58: 211–215.
Nanayama F, Satake K, Furukawa R, et al. (2003) Unusually large earthquakes inferred from tsunami deposits along the Kuril trench. *Nature* 424: 660–663.
Nanayama F, Shigeno K, Satake K, et al. (2000) Sedimentary differences between the 1993 Hokkaido-nansei-oki tsunami and the 1959 Miyakojima typhoon at Taisei, southwestern Hokkaido, northern Japan. *Sedimentary Geology* 135: 255–264.
Neetu S, Suresh I, Shankar R, et al. (2005) Comment on "The Great Sumatra–Andaman earthquake of 26 December 2004" *Science* 310: 1431a.
Nelson AR, Shennan I, and Long AJ (1996) Identifying coseismic subsidence in tidal-wetland stratigraphic sequences at the Cascadia subduction zone of western North America. *Journal of Geophysical Research: Solid Earth* 101: 6115–6135.

Newman AV and Okal EA (1998) Teleseismic estimates of radiated seismic energy: The E/M_0 discriminant for tsunami earthquakes. *Journal of Geophysical Research: Solid Earth* 103: 26885–26898.

Nishenko SP and Buland R (1987) A generic recurrence interval distribution for earthquake forecasting. *Bulletin of the Seismological Society of America* 77: 1382–1399.

Nishimura Y and Miyaji N (1995) Tsunami deposits from the 1993 Southwest Hokkaido earthquake and the 1640 Hokkaido-Komagatake eruption, northern Japan. *Pure and Applied Geophysics* 144: 719–733.

Nomanbhoy N and Satake K (1995) Generation mechanism of tsunamis from the 1883 Krakatau eruption. *Geophysical Research Letters* 22: 509–512.

Okada Y (1985) Surface deformation due to shear and tensile faults in a half-space. *Bulletin of the Seismological Society of America* 75: 1135–1154.

Okada M (1995) Tsunami observation by ocean bottom pressure gauge. In: Tsuchiya Y and Shuto N (eds.) *Tsunami: Progress in Prediction, Disaster Prevention and Warning*, pp. 287–303. Dordrecht: Kluwer Academic Publishers.

Okal EA (1988) Seismic parameters controlling far-field tsunami amplitudes: A review. *Natural Hazards* 1: 67–96.

Okal EA and Newman AV (2001) Tsunami earthquakes: The quest for a regional signal. *Physics of the Earth and Planetary Interiors* 124: 45–70.

Okal EA, Plafker G, Synolakis CE, et al. (2003) Near-field survey of the 1946 Aleutian tsunami on Unimak and Sanak Islands. *Bulletin of the Seismological Society of America* 93: 1226–1234.

Okal EA and Synolakis CE (2004) Source discriminants for near-field tsunamis. *Geophysical Journal International* 158: 899–912.

Okal EA, Synolakis CE, Fryer GJ, et al. (2002) A field survey of the 1946 Aleutian tsunami in the far field. *Seismological Research Letters* 73: 490–503.

Okal EA and Talandier J (1989) M_m: A variable-period mantle magnitude. *Journal of Geophysical Research: Solid Earth* 94: 4169–4193.

Ortiz M and Bilham R (2003) Source area and rupture parameters of the 31 December 1881 $M_w = 7.9$ Car Nicobar earthquake estimated from tsunamis recorded in the Bay of Bengal. *Journal of Geophysical Research: Solid Earth* 108, http://dx.doi.org/10.129/2002JB001941.

Ozaki T (2011) Outline of the 2011 off the Pacific coast of Tohoku earthquake (Mw 9.0)—Tsunami warnings/advisories and observations. *Earth, Planets and Space* 63: 827–830.

Ozawa S, Nishimura T, Suito H, et al. (2011) Coseismic and postseismic slip of the 2011 magnitude-9 Tohoku-oki earthquake. *Nature* 475: 373–376.

Parsons T and Geist E (2008) Tsunami probability in the Caribbean region. *Pure and Applied Geophysics* 165: 2089–2116.

Pelayo AM and Wiens DA (1992) Tsunami earthquakes: Slow thrust-faulting events in the accretionary wedge. *Journal of Geophysical Research: Solid Earth* 97: 15321–15337.

Pelinovsky E, Choi BH, Stromkov A, et al. (2005) Analysis of tide-gauge records of the 1883 Krakatau tsunami. In: Satake K (ed.) *Tsunamis: Case Studies and Recent Developments*, pp. 57–78. Dordrecht: Springer.

Peregrine H (1972) Equations for water waves and the approximations behind them. In: Meyer RE (ed.) *Waves on Beaches and Resulting Sediment Transport*, pp. 95–121. New York: Academic Press.

Piatanesi A and Lorito S (2007) Rupture process of the 2004 Sumatra–Andaman earthquake from tsunami waveform inversion. *Bulletin of the Seismological Society of America* 97: S223–S231.

Piatanesi A, Tinti S, and Gavagni I (1996) The slip distribution of the 1992 Nicaragua earthquake from tsunami run-up data. *Geophysical Research Letters* 23: 37–40.

Pires C and Miranda PMA (2001) Tsunami waveform inversion by adjoint methods. *Journal of Geophysical Research: Oceans* 106: 19773–19796.

Plafker G, Kachadoorian R, Eckel EB, et al. (1969) Effects of the earthquake of March 27, 1964 on various communities. *U. S. G. S. Professional Paper* 542-G, pp. 1–50.

Polet J and Kanamori H (2000) Shallow subduction zone earthquakes and their tsunamigenic potential. *Geophysical Journal International* 142: 684–702.

Prendergast AL, Cupper ML, Jankaew K, et al. (2012) Indian Ocean tsunami recurrence from optical dating of tsunami sand sheets in Thailand. *Marine Geology* 295–298: 20–27.

Press F and Harkrider D (1966) Air-sea waves from the explosion of Krakatoa. *Science* 154: 1325–1327.

Rabinovich AB, Candella RN, and Thomson RE (2011) Energy decay of the 2004 Sumatra tsunami in the world ocean. *Pure and Applied Geophysics* 168: 1919–1950.

Rabinovich AB and Thomson RE (2007) The 26 December 2004 Sumatra tsunami: Analysis of tide gauge data from the world ocean Part 1. Indian Ocean and South Africa. *Pure and Applied Geophysics* 164: 261–308.

Rajendran CP, Rajendran K, Andrade V, et al. (2013) Ages and relative sizes of pre-2004 tsunamis in the Bay of Bengal inferred from geologic evidence in the Andaman and Nicobar Islands. *Journal of Geophysical Research: Solid Earth* 117. http://dx.doi.org/10.1002/JGRB50122.

Rajendran CP, Rejendran K, Machado T, et al. (2006) Evidence of ancient sea surges at the Mamallapuram coast of India and implications for previous Indian Ocean tsunami events. *Current Science* 91: 1242–1247.

Rikitake T (1976) Recurrence of great earthquakes at subduction zones. *Tectonophysics* 35: 335–362.

Rikitake T and Aida I (1988) Tsunami hazard probability in Japan. *Bulletin of the Seismological Society of America* 78: 1268–1278.

Rzadkiewicz SA, Mariotti C, and Heinrich P (1997) Numerical simulation of submarine landslides and their hydraulic effects. *Journal of Waterway, Port, Coastal, and Ocean Engineering* 123: 149–157.

Saito T, Matsuzawa T, Obara K, et al. (2010) Dispersive tsunami of the 2010 Chile earthquake recorded by the high-sampling-rate ocean-bottom pressure gauges. *Geophysical Research Letters* 37: L23303.

Satake K (1988) Effects of bathymetry on tsunami propagation—Application of ray tracing to tsunamis. *Pure and Applied Geophysics* 126: 27–36.

Satake K (1989) Inversion of tsunami waveforms for the estimation of heterogeneous fault motion of large submarine earthquakes—The 1968 Tokachi-Oki and 1983 Japan Sea earthquakes. *Journal of Geophysical Research: Solid Earth* 94: 5627–5636.

Satake K (1993) Depth distribution of coseismic slip along the Nankai Trough, Japan, from joint inversion of geodetic and tsunami data. *Journal of Geophysical Research: Solid Earth* 98: 4553–4565.

Satake K (1994) Mechanism of the 1992 Nicaragua tsunami earthquake. *Geophysical Research Letters* 21: 2519–2522.

Satake K (1995) Linear and nonlinear computations of the 1992 Nicaragua earthquake tsunami. *Pure and Applied Geophysics* 144: 455–470.

Satake K (2007) Volcanic origin of the 1741 Oshima–Oshima tsunami in the Japan Sea. *Earth, Planets and Space* 59: 381–390.

Satake K, Aung TT, Sawai Y, et al. (2006) Tsunami heights and damage along the Myanmar coast from the December 2004 Sumatra–Andaman earthquake. *Earth, Planets and Space* 58: 243–252.

Satake K, Baba T, Hirata K, et al. (2005) Tsunami source of the 2004 off the Kii peninsula earthquakes inferred from offshore tsunami and coastal tide gauges. *Earth, Planets and Space* 57: 173–178.

Satake K, Fujii Y, Harada T, et al. (2013) Time and space distribution of coseismic slip of the 2011 Tohoku earthquake as inferred from tsunami waveform data. *Bulletin of the Seismological Society of America* 103: 1473–1492.

Satake K, Nanayama F, and Yamaki S (2008) Fault models of unusual tsunami in the 17th century along the Kuril trench. *Earth, Planets and Space* 60: 925–935.

Satake K, Nishimura Y, Putra PS, et al. (2013) Tsunami Source of the 2010 Mentawai, Indonesia earthquake inferred from tsunami field survey and waveform modeling. *Pure and Applied Geophysics* 170: 1567–1582.

Satake K, Okada M, and Abe K (1988) Tide gauge response to tsunamis: Measurements at 40 tide gauge stations in Japan. *Journal of Marine Research* 46: 557–571.

Satake K, Shimazaki K, Tsuji Y, et al. (1996) Time and site of a giant earthquake in Cascadia inferred from Japanese tsunami records of January 1700. *Nature* 379: 246–249.

Satake K, Smith JR, and Shinozaki K (2002) Three-dimensional reconstruction and tsunami model of the Nuuanu and Wailau landslides, Hawaii. In: Takahashi E, Lipman P, Garcia M, Naka J, and Aramaki S (eds.) *Hawaiian Volcanoes: Deep Underwater Perspectives*. *Geophysical Monograph*, pp. 333–346. Washington, D.C.: American Geophysical Union.

Satake K and Tanioka Y (1999) Sources of tsunami and tsunamigenic earthquakes in subduction zones. *Pure and Applied Geophysics* 154: 467–483.

Satake K and Tanioka Y (2003) The July 1998 Papua New Guinea earthquake: Mechanism and quantification of unusual tsunami generation. *Pure and Applied Geophysics* 160: 2087–2118.

Satake K, Wang KL, and Atwater BF (2003) Fault slip and seismic moment of the 1700 Cascadia earthquake inferred from Japanese tsunami descriptions. *Journal of Geophysical Research: Solid Earth* 108. http://dx.doi.org/10.1029/2003JB002521.

Sato M, Ishikawa T, Ujihara N, et al. (2011) Displacement above the hypocenter of the 2011 Tohoku-oki earthquake. *Science* 332: 1395.

Sawai Y, Fujii Y, Fujiwara O, et al. (2008) Marine incursions of the past 1500 years and evidence of tsunamis at Suijin-numa, a coastal lake facing the Japan Trench. *Holocene* 18: 517–528.

Sawai Y, Kamataki T, Shishikura M, et al. (2009) Aperiodic recurrence of geologically recorded tsunamis during the past 5500 years in eastern Hokkaido, Japan. *Journal of Geophysical Research: Solid Earth* 114. http://dx.doi.org/10.1029/2007JB005503.

Sawai Y, Namegaya Y, Okamura Y, et al. (2012) Challenges of anticipating the 2011 Tohoku earthquake and tsunami using coastal geology. *Geophysical Research Letters* 39. http://dx.doi.org/10.1029/2012GL053692.

Schulte P, Alegret L, Arenillas I, et al. (2010) The Chicxulub asteroid impact and mass extinction at the Cretaceous–Paleogene boundary. *Science* 327: 1214–1218.
Seno T (2002) Tsunami earthquake as transient phenomena. *Geophysical Research Letters* 29, http://dx.doi.org/10.029/2002GL014868.
Seno T and Hirata K (2007) Did the 2004 Sumatra–Andaman earthquake involve a component of tsunami earthquakes? *Bulletin of the Seismological Society of America* 97: S296–S306.
Shimazaki K, Kim HY, Chiba T, et al. (2011) Geological evidence of recurrent great Kanto earthquakes at the Miura Peninsula, Japan. *Journal of Geophysical Research: Solid Earth* 116. http://dx.doi.org/10.1029/2011JB008639.
Shuto N (1991) Numerical simulation of tsunamis—Its present and near future. *Natural Hazards* 4: 171–191.
Shuto N and Matsutomi H (1995) Field survey of the 1993 Hokkaido-nansei-oki earthquake tsunami. *Pure and Applied Geophysics* 144: 649–663.
Simkin T and Fiske R (1983) *Krakatau, 1883—The Volcanic Eruption and Its Effects*, 464 pp. Washington, D.C.: Smithsonian Institution Press.
Singh SK, Ortiz M, Gupta HK, et al. (2006) Slow slip below Port Blair, Andaman, during the great Sumatra–Andaman earthquake of 26 December 2004. *Geophysical Research Letters* 33. http://dx.doi.org/10.1029/2005GL025025.
Slingerland R and Voight B (1979) Occurrence, properties, and predictive models of landslide-generated water waves. In: Voight B (ed.) *Rockslides and Avalanches*, pp. 317–397. Amsterdam: Elsevier.
Smith WHF, Scharroo R, Titov VV, et al. (2005) Satellite altimeters measure tsunami, early model estimates confirmed. *Oceanography* 18: 10–12.
Smith DE, Shi S, Cullingford RA, et al. (2004) The Holocene Storegga slide tsunami in the United Kingdom. *Quaternary Science Reviews* 23: 2291–2321.
Soloviev SL (1970) Recurrence of tsunamis in the Pacific. In: Adams WM (ed.) *Tsunamis in the Pacific Ocean*, pp. 149–164. Honolulu: East–West Center Press.
Soloviev SL and Go CN (1974a) *Catalogue of Tsunamis on the Eastern Shore of the Pacific Ocean (1513–1968)*, 204 pp. Moscow: Nauka Publishing House.
Soloviev SL and Go CN (1974b) *Catalogue of Tsunamis on the Western Shore of the Pacific Ocean (173–1968)*, 310 pp. Moscow: Nauka Publishing House.
Soloviev SL, Go CN, and Kim KS (1992) *Catalog of Tsunamis in the Pacific 1969–1982*, 208 pp. Moscow: Academy of Science of the USSR.
Song YT, Ji C, Fu L-L, et al. (2005) The 26 December 2004 tsunami source estimated from satellite radar altimetry and seismic waves. *Geophysical Research Letters* 32. http://dx.doi.org/10.1029/2005GL023683.
Stein S and Okal EA (2005) Speed and size of the Sumatra earthquake. *Nature* 434: 581–582.
Steketee JA (1958) On Volterra's dislocations in a semi-infinite elastic medium. *Canadian Journal of Physics* 36: 192–205.
Sugawara D, Imamura F, Goto K, et al. (2013) The 2011 Tohoku-oki earthquake tsunami: Similarities and differences to the 869 Jogan tsunami on the Sendai Plain. *Pure and Applied Geophysics* 170: 831–843.
Synolakis CE, Bardet JP, Borrero JC, et al. (2002) The slump origin of the 1998 Papua New Guinea tsunami. *Proceedings of the Royal Society of London, Series A* 458: 763–789.
Synolakis CE and Okal EA (2005) 1992–2002: Perspective on a decade of post-tsunami surveys. In: Satake K (ed.) *Tsunamis: Case Studies and Recent Developments*, pp. 1–29. Dordrecht: Springer.
Takahashi T, Shuto N, Imamura F, et al. (1995) Source models for the 1993 Hokkaido-nansei-oki earthquake tsunami. *Pure and Applied Geophysics* 144: 747–767.
Tang L, Titov VV, Bernard EN, et al. (2012) Direct energy estimation of the 2011 Japan tsunami using deep-ocean pressure measurements. *Journal of Geophysical Research: Oceans* 117: C08008. http://dx.doi.org/10.1029/2011JC007635.
Tang L, Titov VV, and Chamberlin CD (2009) Development, testing, and applications of site-specific tsunami inundation models for real-time forecasting. *Journal of Geophysical Research: Oceans* 114. http://dx.doi.org/10.1029/2009JC005476.
Tanioka Y, Nishimura Y, Hirakawa K, et al. (2004) Tsunami run-up heights of the 2003 Tokachi-oki earthquake. *Earth, Planets and Space* 56: 359–365.
Tanioka Y, Ruff L, and Satake K (1997) What controls the lateral variation of large earthquake occurrence along the Japan trench? *Island Arc* 6: 261–266.
Tanioka Y and Satake K (1996a) Fault parameters of the 1896 Sanriku tsunami earthquake estimated from tsunami numerical modeling. *Geophysical Research Letters* 23: 1549–1552.
Tanioka Y and Satake K (1996b) Tsunami generation by horizontal displacement of ocean bottom. *Geophysical Research Letters* 23: 861–864.
Tanioka Y and Satake K (2001a) Coseismic slip distribution of the 1946 Nankai earthquake and aseismic slips caused by the earthquake. *Earth, Planets and Space* 53: 235–241.
Tanioka Y and Satake K (2001b) Detailed coseismic slip distribution of the 1944 Tonankai earthquake estimated from tsunami waveforms. *Geophysical Research Letters* 28: 1075–1078.
Tanioka Y, Satake K, and Ruff L (1995) Total analysis of the 1993 Hokkaido-nansei-oki earthquake using seismic-wave, tsunami, and geodetic data. *Geophysical Research Letters* 22: 9–12.
Tanioka Y and Seno T (2001a) Detailed analysis of tsunami waveforms generated by the 1946 Aleutian tsunami earthquake. *Natural Hazards and Earth System Sciences* 1: 171–175.
Tanioka Y and Seno T (2001b) Sediment effect on tsunami generation of the 1896 Sanriku tsunami earthquake. *Geophysical Research Letters* 28: 3389–3392.
Tanioka Y, Yudhicara, Kusunose T, et al. (2006) Rupture process of the 2004 great Sumatra–Andaman earthquake estimated from tsunami waveforms. *Earth, Planets and Space* 58: 203–209.
Tappin DR (2004) Special issue: Submarine-slump-generated tsunamis—Preface. *Marine Geology* 203: 199–200.
Tappin DR, Matsumoto T, Watts P, et al. (1999) Sediment slump likely caused 1998 Papua New Guinea tsunami. *Eos, Transactions American Geophysical Union* 80: 329, 334, 340.
Tatehata H (1997) The new tsunami warning system of the Japan Meteorological Agency. In: Hebenstreit G (ed.) *Perspectives on Tsunami Hazards Reduction*, pp. 175–188. Dordrecht: Kluwer Academic Publishers.
Tilling RI (1989) Volcanic hazards and their mitigation—Progress and problems. *Reviews of Geophysics* 27: 237–269.
Tinti S, Bortolucci E, and Vannini C (1997) A block-based theoretical model suited to gravitational sliding. *Natural Hazards* 16: 1–28.
Tinti S, Maramai A, Armigliato A, et al. (2005) Observations of physical effects from tsunamis of December 30, 2002 at Stromboli volcano, southern Italy. *Bulletin of Volcanology* 68: 450–461.
Titov VV and Gonzalez FI (2001) Numerical study of the source of the July 17, 1998 PNG tsunami. In: Hebensreit GT (ed.) *Tsunami Research at the End of a Critical Decade*, pp. 197–207. Dordrecht: Kluwer Acad. Pub.
Titov VV, Gonzalez FI, Bernard EN, et al. (2005) Real-time tsunami forecasting: Challenges and solutions. *Natural Hazards* 35: 41–58.
Titov V, Rabinovich AB, Mofjeld HO, et al. (2005) The global reach of the 26 December 2004 Sumatra tsunami. *Science* 309: 2045–2048.
Titov VV and Synolakis CE (1997) Extreme inundation flows during the Hokkaido-nansei-oki tsunami. *Geophysical Research Letters* 24: 1315–1318.
Tobita M, Suito H, Imakiire T, et al. (2006) Outline of vertical displacement of the 2004 and 2005 Sumatra earthquakes revealed by satellite radar imagery. *Earth, Planets and Space* 58: e1–e4.
Todorovska MI, Hayir A, and Trifunac MD (2003) Near-field amplitudes of tsunami from submarine slumps and slides. In: Yalçıner AC, Pelinovsky E, Okal E, and Synolakis CE (eds.) *Submarine Landslides and Tsunamis*, pp. 59–68. The Netherlands: Kluwer.
Toh H, Satake K, Hamano Y, et al. (2011) Tsunami signals from the 2006 and 2007 Kuril earthquakes detected at a seafloor geomagnetic observatory. *Journal of Geophysical Research: Solid Earth* 116. http://dx.doi.org/10.1029/2010JB007873.
Tsai VC, Ampuero J-P, Kanamori H, et al. (2013) Estimating the effect of Earth elasticity and variable water density on tsunami speeds. *Geophysical Research Letters* 40. http://dx.doi.org/10.1002/grl.50147.
Tsai VC, Nettles M, Ekstrom G, et al. (2005) Multiple CMT source analysis of the 2004 Sumatra earthquake. *Geophysical Research Letters* 32. http://dx.doi.org/10.1029/2005GL023813.
Tsuboi S, Abe K, Takano K, et al. (1995) Rapid-determination of Mw from broad-band P-wave-forms. *Bulletin of the Seismological Society of America* 85: 606–613.
Tsuji Y, Namegaya Y, Matsumoto H, et al. (2006) The 2004 Indian tsunami in Thailand: Surveyed run-up heights and tide gauge records. *Earth, Planets and Space* 58: 223–232.
Tuttle MP, Ruffman A, Anderson T, et al. (2004) Distinguishing tsunami from storm deposits in eastern North America: The 1929 grand banks tsunami versus the 1991 Halloween storm. *Seismological Research Letters* 75: 117–131.
Unoki S and Nakano M (1953) On the Cauchy–Poisson waves caused by the eruption of a submarine volcano (III). *Oceanographical Magazine* 4: 139–150.
Utsu T (1984) Estimation of parameters for recurrence models of earthquakes. *Bulletin of the Earthquake Research Institute, University of Tokyo* 59: 53–66.
Ward S (1980) Relationships of tsunami generation and an earthquake source. *Journal of Physics of the Earth* 28: 441–474.
Ward SN (2001) Landslide tsunami. *Journal of Geophysical Research: Solid Earth* 106: 11201–11215.
Ward SN and Day S (2003) Ritter Island Volcano—Lateral collapse and the tsunami of 1888. *Geophysical Journal International* 154: 891–902.
Watada S, Kusumoto S, and Satake K (2013) Cause of travel-time difference between observed and synthetic waveforms distant tsunami (Abstract). In: *IAHS-IASPO-IASPEI Meeting*.

Watanabe H (1998) *Comprehensive List of Tsunamis to Hit the Japanese Islands (in Japanese)*, 238 pp. Tokyo: University of Tokyo Press.

Watts P, Grilli ST, Kirby JT, et al. (2003) Landslide tsunami case studies using a Boussinesq model and a fully nonlinear tsunami generation model. *Natural Hazards and Earth System Sciences* 3: 391–402.

Wei Y, Bernard EN, Tang L, et al. (2008) Real-time experimental forecast of the Peruvian tsunami of August 2007 for U.S. coastlines. *Geophysical Research Letters* 35. http://dx.doi.org/10.1029/2007GL032250.

Yamashita T and Sato R (1974) Generation of tsunami by a fault model. *Journal of Physics of the Earth* 22: 415–440.

Yamazaki Y and Cheung KF (2011) Shelf resonance and impact of near-field tsunami generated by the 2010 Chile earthquake. *Geophysical Research Letters* 38: L12605.

Yeh H, Imamura F, Synolakis C, et al. (1993) The Flores Island tsunamis. *Eos, Transactions American Geophysical Union* 74: 369, 371–373.

Yokoyama I (1987) A scenario of the 1883 Krakatau tsunami. *Journal of Volcanology and Geothermal Research* 34: 123–132.

Yuan XH, Kind R, and Pedersen HA (2005) Seismic monitoring of the Indian Ocean tsunami. *Geophysical Research Letters* 32. http://dx.doi.org/10.1029/2005GL023464.

4.20 Physical Processes That Control Strong Ground Motion

JG Anderson, University of Nevada, Reno, NV, USA

© 2015 Elsevier B.V. All rights reserved.

4.20.1	**Introduction**	505
4.20.2	**What Is Strong Ground Motion?**	506
4.20.2.1	Strong-Motion Instrumentation	506
4.20.2.2	Parameters to Characterize Strong Motion	508
4.20.2.3	Amplitudes of Strong Motions	508
4.20.3	**Representation Theorem**	511
4.20.4	**Green's Functions**	514
4.20.4.1	Synthetic Green's Functions	514
4.20.4.1.1	Point source in an unbounded medium	514
4.20.4.1.2	Green's functions for more realistic velocity models	515
4.20.4.2	Empirical Green's Functions	515
4.20.4.2.1	Effect of signal-to-noise ratio for EGFs	516
4.20.5	**Source Parameters**	516
4.20.5.1	The Haskell Model	516
4.20.5.2	Example of a Slip Model	517
4.20.5.3	Fourier Amplitude Spectrum of the Seismic Source	518
4.20.5.4	Scaling of Source Parameters	519
4.20.6	**The Effects of Fault Finiteness**	521
4.20.6.1	Spectral Shape	521
4.20.6.2	Directivity and Duration of Shaking	521
4.20.7	**Effects of Surface Geology**	522
4.20.7.1	NEHRP Site Classifications	523
4.20.7.2	Amplification Due to the Decrease in Shear Velocity	524
4.20.7.3	Refraction of Body Waves Toward Vertical Incidence	525
4.20.7.4	Reflection and Transmission Across Interfaces	525
4.20.7.5	Amplification and Resonance in Flat, Low-Velocity Layers at the Surface	526
4.20.7.6	Trapping Energy in the Form of Surface Waves	527
4.20.7.7	Effect of Basin Edges on Incoming Body Waves	527
4.20.7.8	Trapping and Resonance in Basins	528
4.20.7.9	Long Linear Valleys Acting as Waveguides	529
4.20.7.10	Multipathing	530
4.20.7.11	Scattering	530
4.20.7.12	Correlation and Coherency	531
4.20.7.13	Topography	532
4.20.7.14	Q	533
4.20.7.15	Summary: Spectral Model	535
4.20.7.16	Nonlinearity	535
4.20.7.17	Seismic Site–City Interactions	540
4.20.8	**Predicting Strong Ground Motions**	540
4.20.8.1	Synthetic Seismograms	540
4.20.8.1.1	Stochastic simulations	540
4.20.8.1.2	Empirical Green's functions	541
4.20.8.1.3	Synthetic Green's functions with kinematic and dynamic sources	541
4.20.8.2	Ground Motion Prediction Equations	542
4.20.9	**Conclusions**	547
Acknowledgments		548
References		548

4.20.1 Introduction

Strong-motion seismology is the branch of seismology specializing in understanding ground motions near large earthquakes. It is important first for the practical expectation of our civilization that structures will be built to withstand earthquakes, and this cannot be done without knowing the character of those ground motions. The second reason it is important is that the ground motions recorded near the earthquake are generally least filtered and modified by transmission through the

 http://dx.doi.org/10.1016/B978-0-444-53802-4.00087-7

Earth and thus give the best data for scientific study of the earthquake source.

The goal of strong-motion seismology is to be able to predict motions near an earthquake sufficiently well that the predictions can be used for engineering applications, for both scientific and practical reasons. The scientific motivation is that the ability to predict the motions is evidence that the phenomena are understood. The practical reason is the need to know the character of ground motions at a site of engineering interest, given that the sources and sites, and the geologic features between, are unique to every situation. In most cases, the character of strong ground motions is approximately understandable, and the amplitudes of the motions critical for engineering designs are predictable within modest uncertainties.

Some recent relatively extensive reviews of aspects of research in the field of strong-motion seismology include Takemura et al. (1995) and Anderson (2003). The volume compiled by Lee et al. (2003) includes several articles that give detailed reviews of aspects of the field. While there is overlap of topics considered here and in those other recent reviews, the emphasis in this chapter is on the physical processes that affect strong ground motions.

4.20.2 What Is Strong Ground Motion?

Strong ground motion is an event in which an earthquake shakes the ground at least strongly enough for people to feel the motion or, worse, strongly enough to damage or destroy man-made structures.

The field of strong-motion seismology could initially be identified with a type of instrument, designed to remain on-scale and record the ground motion with fidelity under the conditions of the strongest ground motions experienced in earthquakes. Special instrumentation was necessary because the amplitudes of strong motion are orders of magnitude greater than motions recorded from teleseisms or small local earthquakes. Sensitive seismic instruments optimized to maximize the number of earthquakes that are recorded are driven off scale by ground motions sufficient to cause damage.

4.20.2.1 Strong-Motion Instrumentation

The displacement of the mass of a mechanical seismometer is proportional to the acceleration of the ground when the natural frequency of the system is greater than the frequency of the ground motions. A high natural frequency, achieved with a small mass and stiff spring, is preferable to achieve low sensitivity for strong-motion studies. Instruments providing seismogram traces on which the deflections are proportional to acceleration are called accelerographs. With acceleration being the property that is most easily read from strong-motion records and with forces on structures proportional to acceleration, peak acceleration of the ground became the most commonly cited measure of strong ground motions. Commonly used units in strong-motion seismology are 'g,' the reference acceleration of gravity on Earth, and the Gal (named for Galileo), where $1\ \text{Gal} = 1\ \text{cm s}^{-2}$ and $980.665\ \text{Gal} = 9.80665\ \text{ms}^{-1} = 1\ \text{g}$.

The history of instrumentation in the United States is reviewed by Trifunac and Todorovska (2001). The first US instruments were deployed in 1932 by the US Coast and Geodetic Survey (**Figure 1(a)**), and the first significant US record was obtained in 10 March 1933 Long Beach, California, earthquake (M_w 6.4). The recorded shaking in Long Beach, California, during the 1933 Long Beach earthquake had peak horizontal and vertical accelerations of 0.19 and 0.28 g, respectively. The history of the Japanese program, reviewed by Takemura et al. (1995), began in 1951 (**Figure 1(b)**) and the first significant Japanese record was obtained in 1956 (Takahasi, 1956). Early instruments were typically designed so that ground motions up to the acceleration of gravity (1 g) would be on-scale. Going off scale meant that a light beam moved off the edge of the recording film (**Figure 1(a)** and **1(c)**) or that a pen scratching a waxed paper (**Figure 1(b)**) swung off the edge of the paper. In one widely used compact photographic accelerograph (**Figure 1(c)**), while the instrument was at rest, the light beam would be at least 2 cm from the edge of the film, and the instrument was designed such that a 2 cm deflection occurred under 1 g acceleration. The thickness of the light beam defined the minimum ground motion that could be recorded. Supposing that a well-focused light beam exposes the film in a line no more than 0.5 mm wide and that the center of the beam can be estimated to 0.2 mm, the minimum acceleration resolved is somewhat less than 0.01 g. This roughly defined the lower limit of ground motion considered by the early strong-motion seismology studies. Conveniently, 0.01 g is only a little stronger than the lower limit of ground motions that a person at rest is able to feel. Thus, the early range of strong-motion seismology approximately coincided with ground motions that humans are able to feel.

Modern strong-motion instruments (e.g., **Figure 1(d)**) use digital recording with at least 16-bit and typically 24-bit analog-to-digital conversion (e.g., Kinoshita, 2003; Shin et al., 2003). With a 24-bit instrument, assuming conservatively that there are at least 19 bits above noise sources, the least significant bit in a 1 g accelerograph is reduced to under 4×10^{-6} g, or $<0.004\ \text{cm s}^{-2}$. This increase in the dynamic range of the recording system takes advantage of force-balance sensors, described, for example, by Wielandt (2002). Where the dynamic range of the analog instrument was about 40 dB ($=20 \log(1\,\text{g}/0.01)$), the dynamic range of the current generation of digital accelerographs is between 110 and 135 dB (Trifunac and Todorovska, 2001). Since much smaller ground motions can be recorded on these instruments than on the earlier generations of analog instruments, the distinction between strong-motion seismology and traditional seismology is blurred. While analog seismogram recordings are now considered of limited use for most applications in seismology, as noted by Scherbaum (2002), strong-motion accelerograms from significant earthquakes recorded on analog instruments are still very valuable because strong motions are rare events. Because it is rare, collections of strong-motion data are available on the World Wide Web. **Appendix** lists the websites for several selected strong-motion programs throughout the world.

The KiK-net strong-motion stations record acceleration with sufficient precision that after the 2011 Tōhoku-Oki earthquake, they were able to be integrated to recover static displacement of the ground (Kinoshita and Takagishi, 2011). On the other hand, the location using the Global Positioning System (GPS) is now also capable of recovering ground displacements at a sample rate of 1 Hz

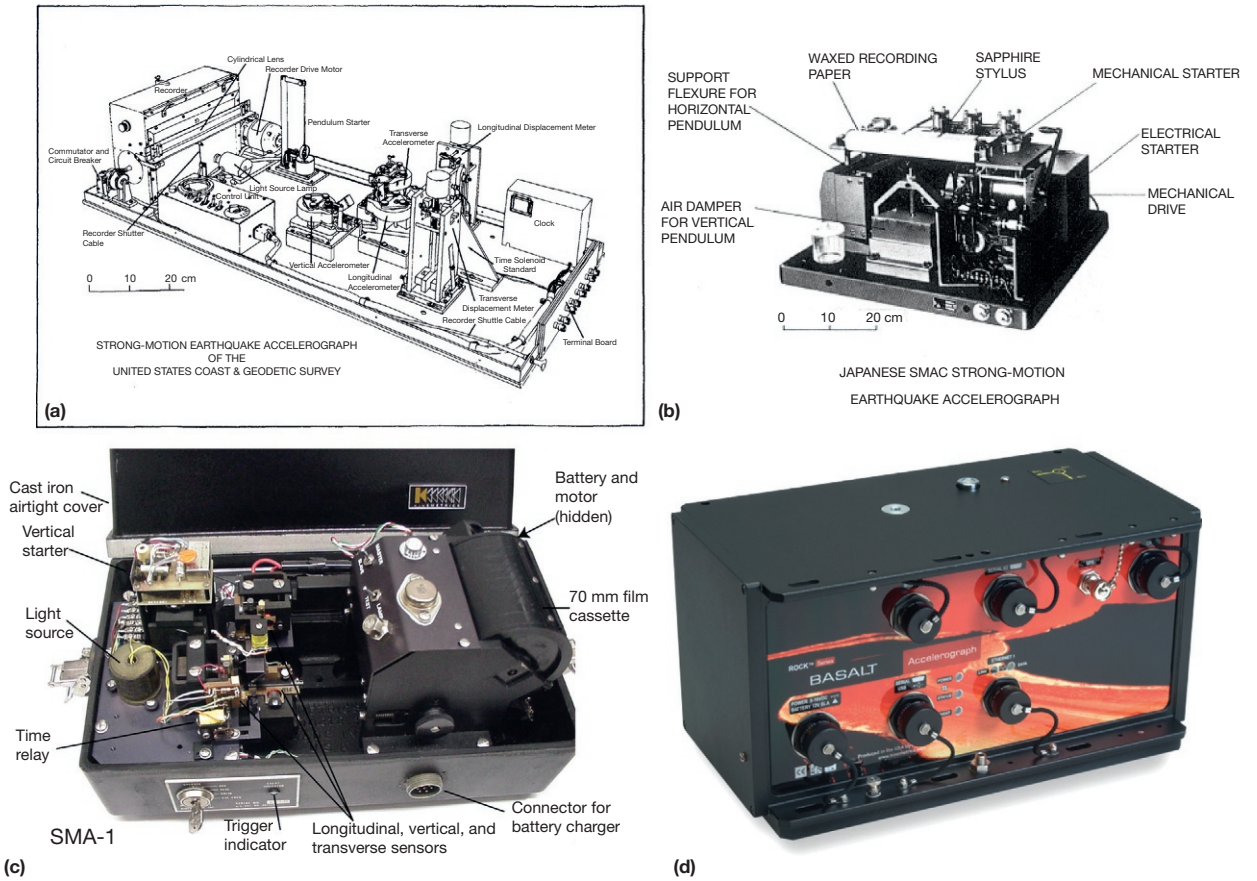


Figure 1 Several strong-motion accelerographs. (a) US Coast and Geodetic Survey accelerograph, deployed from the 1932s to the 1960s. Hudson (1963) reported that at that time, there were 60 in operation in the United States. Illustration from Hudson (1963). (b) SMAC accelerograph, deployed in Japan from the early 1950s to the 1960s. Hudson (1963) reported that at that time, there were 50 in operation in Japan and 15 accelerographs of another design. Illustration from Hudson (1963). (c) SMA-1 (strong-motion accelerograph). This compact strong-motion accelerograph, manufactured by Kinemetrics, Inc., was widely deployed in the 1970s and 1980s, with over 6600 sold. The iron case was designed to prevent damage to the internal instrument in the case of objects falling on the instrument during strong motion. Internal batteries were trickle-charged so that the instrument would not depend on power during the earthquake. The instrument was triggered by a separate vertical sensor and, after the last triggering acceleration, was set to run for, typically, 10 s. After triggering, the instrument started recording in about 0.1 s, and three components of acceleration were recorded on 70 mm wide film transported through the camera at a rate of 1 cm s^{-1}. Time relays place 0.5 s marks or WWVB time code along the edge of the film. A single roll of film provided for about 15 min of recording time. The sensors consisted of mirrors mounted on leaf springs with natural frequency of about 20 Hz. Light beams reflected off the mirrors were recorded on the film. (d) Basalt strong-motion accelerograph, manufactured by Kinemetrics, Inc., in the early 2010s. Sensors are force-balance accelerometers, from which acceleration is digitized with a 24-bit analog-to-digital converter at up to 2000 samples per second; 200 samples per second is typical. Time is provided by an internal clock synchronized to the Global Positioning System. Output from the accelerometers is continuously monitored, and when triggering motions occur, the accelerogram including pre-event memory is saved on internal compact flash memory. Data also can be sent immediately to a central location with a choice of telemetry alternatives. Instrument health can be monitored, data can be retrieved, and settings can be modified through access over the World Wide Web.

(e.g., Branzanti et al., 2013; Genrich and Bock, 2006; Larson et al., 2003). Between these two data sources, ground motions over an unprecedented broad bandwidth can now be measured. Uses include tsunami warning, rapid magnitude determination, and broadband source inversions in addition to the current uses of immediate risk assessment and early warning.

Improvements in instrumentation were accompanied by an increased determination by the earthquake engineering and seismological communities to expand the numbers of strong-motion instruments. Hudson (1963) identified two major strong-motion networks at that time: about 60 USCGS instruments in the United States and 65 instruments in Japan. The 1960s saw the beginning of a significant expansion (Trifunac and Todorovska, 2001). The International Workshop on Strong-Motion Instrumentation Arrays in 1978 was influential in motivating expanded instrumentation focused on locations with high probabilities of recording strong earthquakes (Iwan, 1979). Some significant earthquakes have motivated governments to respond to the need, articulated by their earthquake professionals, for expanded instrumentation. For instance, the 1995 Kobe, Japan, earthquake was followed by the deployment of the K-NET and KiK-net strong-motion networks in Japan (Kinoshita, 2003; Shin et al., 2003).

4.20.2.2 Parameters to Characterize Strong Motion

Because strong-motion records are critical for the safe design of structures, considerable care is taken to assure that the instruments are well calibrated and the accelerograms are processed carefully. The processing of strong-motion accelerograms is reviewed by, among others, Takemura et al. (1995), Shakal et al. (2003), and Boore and Bommer (2005). Processing includes making a 'baseline correction' because a raw accelerogram has a nonzero average value due to imperfect leveling of the instrument. For analog instrument recording on film, there is no indication of the location of 'zero' acceleration on the film. During severe shaking, the instrument may be tilted slightly compared to its position prior to the earthquake, complicating baseline identification. There is always some level of low- and high-frequency noise in even the best digital records, so processing generally involves band-pass filtering to select only those frequencies with an adequate signal-to-noise ratio. An instrument correction is applied. After these steps, the acceleration can be integrated to obtain velocity and displacement, used to compute Fourier and response spectra, and used to characterize the effects on structures (e.g., see Hall, 2003; Jennings, 2003). Low-frequency pulses that carry a permanent displacement of the ground associated with slip on the fault are present in some cases of strong ground motion. The static displacement has been recovered in a few cases by double integration of digital strong-motion accelerograms with sufficient resolution (e.g., Anderson et al., 1986, 2000b; Boore et al., 2002; Chanerley et al., 2013; Ma et al., 2001). Javelaud et al. (2011) suggesting a systematic approach to baseline correction in cases where a static offset is present in the data.

Strong ground motion is complex, so it cannot be characterized with one or even a small number of parameters. Since it is recorded with an accelerometer, the most common parameter used to characterize a record is the peak acceleration. The peak velocity and, less often, peak displacement are also used to characterize a record. The duration of strong shaking can be measured in two ways: One definition is the time during which most of the energy (e.g., 90%) arrives (e.g., Trifunac and Brady, 1975) and the other is the time between the first peak and the last peak exceeding some threshold (e.g., 5%g, Bolt, 1969; Kramer, 1996). As distance from the source increases, amplitudes decrease and waves progressively disperse, so that the duration by the first definition increases as distance increases, while the duration by the second definition decreases to zero (no strong shaking if the selected threshold is not exceeded). Two classes of spectra are also important: First, Fourier spectra are used to decompose the shaking into contributions from waves as a function of frequency and to make inferences about the source. Second, the response spectra measure the peak response (displacement, velocity, or acceleration) of the mass of an idealized, damped (usually at 5% of critical damping) single-degree-of-freedom oscillator when the base of the oscillator is subjected to the accelerogram (e.g., Hudson, 1979). Response spectra are formed by synthesizing the response of a suite of damped oscillators with different natural periods. There are five types of response spectra: displacement response ($SD(T_n,\zeta)$), velocity response ($SV(T_n,\zeta)$), pseudo-velocity response ($PSV(T_n,\zeta)=(2\pi/T_n)SD(T_n,\zeta)$), absolute acceleration response ($SA(T_n,\zeta)$), and pseudo-absolute acceleration response ($PSA(T_n,\zeta)=(2\pi/T_n)^2 SD(T_n,\zeta)$), where T_n is the undamped natural period of the oscillator and ζ is the fraction of critical damping. The response spectrum is perhaps the most basic tool and the most used measure of strong motion for engineering applications (Jennings, 2003). It forms the basis for establishing a design spectrum, which specifies the required strength of structures constructed under building codes and structures of major importance where a separate seismic hazard analysis is performed.

4.20.2.3 Amplitudes of Strong Motions

Because motion of the ground during an earthquake is a three-component vector, it is important to designate how that vector is reduced to a scalar in a simple description of 'peak acceleration.' Many approaches have been used (e.g., Boore, 2010; Boore et al., 2006; Kramer, 1996). Because the orientation of the horizontal components of the accelerogram is 'random' relative to the source, parameters that are independent of instrument orientation (e.g., Boore, 2010; Boore et al., 2006) are preferable to parameters on a component oriented as recorded.

Figures 2 and **3** show the properties of the horizontal motions based on an extension of Anderson (2010, 2012). **Figure 2** considers the maximum amplitudes of horizontal vector acceleration (A_H) and velocity (V_H) in a collection of 445 records of strong motion, selected in a way intended to include as many as possible of the strongest records that are available publicly. Only records from 'free-field' or small structure conditions are included in **Figure 2**. Records from large structures may be deamplified if recorded in the basement, due to the inertial mass of the structure, or amplified by the structure if recorded on the roof, due to the flexibility of the structure as it responds to the ground shaking.

Figure 2(a) shows the magnitudes of earthquakes contributing to these strong records, shown by the time the earthquake occurred. The increase in the rate of contributing earthquakes is a direct function of the increase in strong-motion instrumentation since the first strong-motion accelerometers were installed in the early 1930s. **Figure 2(b)** shows A_H in this collection of records as a function of magnitude, and **Figure 2(c)** shows the same for V_H. The largest earthquakes in this collection are in subduction zones, and thus, the stations are all relatively far from the fault. The current empirical upper bounds on both A_H and V_H as a function of magnitude are dominated by records obtained close to earthquakes in continental crust. The current empirical upper bound of V_H increases more strongly as a function of magnitude than the current empirical upper bound on A_H. **Figure 2(d)** shows the ranking of A_H in this collection of accelerograms. There are 71 records with A_H greater than 1 g. **Figure 2(e)** similarly shows 41 records with V_H greater than 100 cm s^{-1}. The four strongest horizontal accelerograms in this compilation exceed 2 g, and the five strongest velocities exceed 1.5 m s^{-1}. Records with such large amplitudes are very rare. Anderson (2010) suggested the terminology 'exceptional ground motions' to apply to records with peak acceleration in excess of 750 cm s^{-2} or velocities in excess of 75 cm s^{-1} and estimated that these amplitudes occur at rates of roughly 7×10^{-4} per instrument year in the strong-motion data, based on an estimate that the accelerogram collection

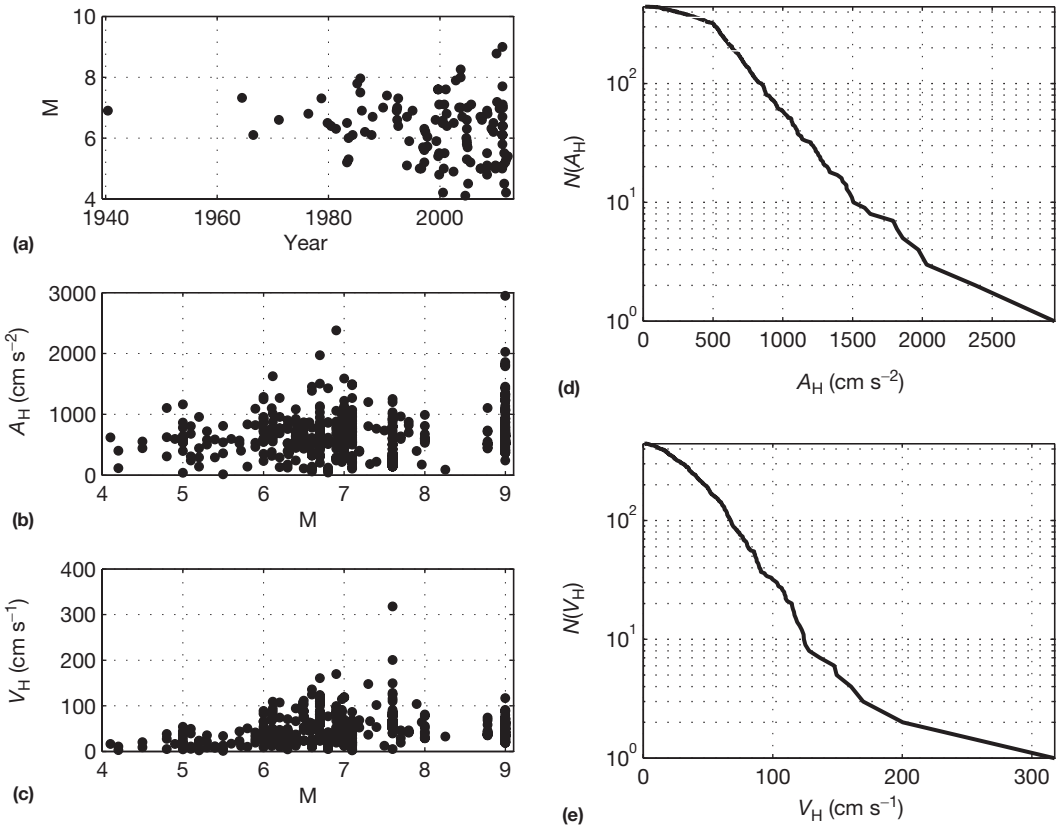


Figure 2 Some characteristics of recordings of strong ground motions that have yielded the largest observed values in the 'free field' or in small structures, extended from Anderson (2010, 2012) to include more recent records. (a) Magnitude of contributing earthquakes as a function of time. (b) Peak horizontal acceleration, A_H, and (c) peak horizontal velocity, V_H, as a function of magnitude. Peak values are the maximum amplitude reached during the event of the horizontal vector formed from the perpendicular horizontal components of the strong-motion recordings. (d) $N(A_H)$, the number of times A_H is equaled or exceeded among the strong-motion records included in this compilation of accelerograms. (e) $N(V_H)$, the number of times V_H is equaled or exceeded.

represented about 150 000 instrument years of observations. Figure 3 shows the pseudo-acceleration response spectra (PSA_H) for each of the accelerograms in Figure 3.

Anderson (2010) reported that about 75% of the records on an earlier version of Figure 2 come from strike-slip, thrust, or oblique thrust-faulting earthquakes in continental crust, while 20% are associated with subduction processes. That ratio changed with the 2011 Tōhoku-Oki, Japan, earthquake (M_W 9.0). While ground motions on rock were not exceptionally large except for their durations (Anderson, 2013), the global number of records with peak accelerations exceeding 1 g was nearly doubled due to amplification of the incoming waves by thin low-velocity layers. Earthquakes with normal-faulting mechanisms were very poorly represented in the database, so it is not possible to conclude from these data that motion resulting from normal-faulting mechanisms is inherently less efficient at generating exceptional ground motions. Most points on Figure 2 are from earthquakes in the magnitude 6.5–7.0 range. Hanks and Johnson (1976) suggested that extremes in peak acceleration should be independent of magnitude at 10 km distance from the fault, where the distance is generalizable to any distance close to the fault. Their hypothesis is that this should be true if the dynamic shear stresses associated with localized failure on the fault are independent of magnitude. Subsequent considerations of the increase in bandwidth associated with increased magnitude found that a weak magnitude dependence should be expected for average peak accelerations (e.g., Baltay et al., 2013; Hanks, 1979; Hanks and McGuire, 1981; McGuire and Hanks, 1980). Figure 2 is consistent with this model.

Figure 2(e) shows 41 records for which the geometric mean of the two horizontal components exceeded 100 cm s^{-1} and about 100 records in excess of 75 cm s^{-1}. A preliminary conclusion, suggested by Anderson et al. (2005a) and Anderson (2010), is that most of the extreme velocities are associated with near-fault pulses for earthquakes with $M > 6$, as will be discussed later in this chapter. The largest peak velocities come from the largest crustal earthquakes in the data.

Tables 1–4 list the top five records sorted by peak vector horizontal acceleration, peak vertical acceleration, peak vector horizontal velocity, and peak vertical velocity. Figure 4 shows the MYG004 accelerogram, which is the strongest peak horizontal acceleration in Table 1. As expected, this record sets the empirical upper bound for most of the high frequencies. The accelerogram itself shows bursts of particularly intense motions from two distinct strong-motion generating areas (Asano and Iwata, 2011; Ide et al., 2011; Kurihashi and Irikura, 2011; Miyake et al., 2003), with the high peak caused by a particularly strong pulse from the second of the

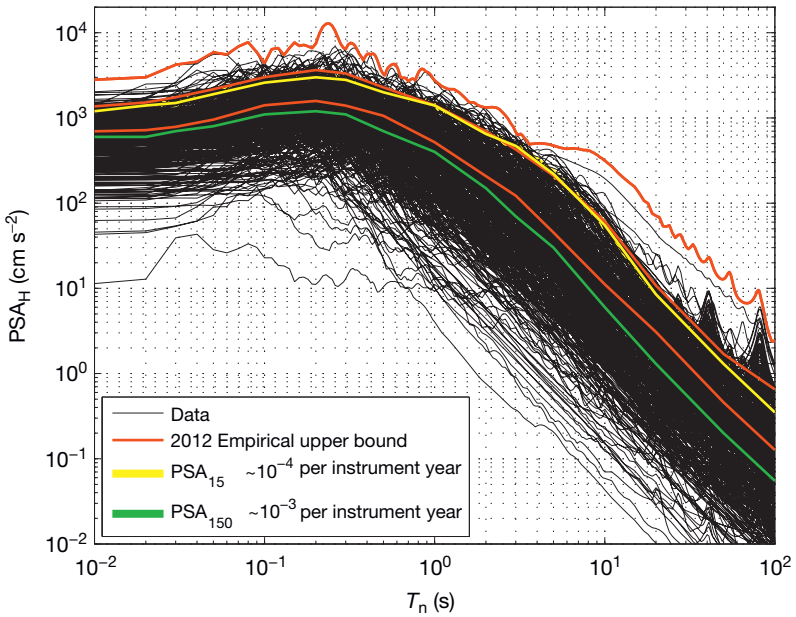


Figure 3 Pseudo-acceleration response spectra corresponding to the 445 accelerograms in **Figure 1**. For each instrument period, T_n, the response of the oscillator to both horizontal components is calculated, the vector amplitude is calculated, and the peak vector amplitude is used to determine the response spectrum, PSA_H (T_n, ζ). Spectra that deal with the vector nature of the horizontal ground motion are called RotD100 spectra (Boore, 2010). The response direction depends on the period. The yellow and green spectral amplitudes are estimated from the smaller data collection by Anderson (2010).

Table 1 Lists of the highest horizontal vector accelerations

Date	*Eq. name*	M_W	*St. name*	A_H (*cm* s^{-2})
11 March 2011	Tōhoku	8.1	MYG004	2762
13 June 2008	Iwate–Miyagi Nairiku	6.9	AKTH04	2504
11 March 2011	Tōhoku	8.1	MYG012	1985
17 January 1994	Northridge	6.7	Tarzana	1965
11 March 2011	Tōhoku	8.1	IBR003	1859

Table 2 Lists of the highest vertical vector accelerations

Date	*Eq. name*	M_W	*St. name*	A_Z (*cm* s^{-2})
13 June 2008	Iwate–Miyagke Nairiku	6.9	IWTH25	3842
23 December 1985	Nahanni	6.9	Site 1	2319
11 March 2011	Tōhoku	9.1	MYG004	1882
15 October 1979	Imperial Valley	6.5	El Centro Array #6	1773
21 February 2011	Christchurch	6.2	NZPRPZ	1598

Table 3 Lists of the highest horizontal vector velocities

Date	*Eq. name*	M_W	*St. name*	V_H (*cm* s^{-1})
20 September 1999	Chi-Chi	7.6	TCU068	306
20 September 1999	Chi-Chi	7.6	TCU052	196
20 September 1999	Chi-Chi	7.6	CHY080	169
17 January 1994	Northridge	6.7	Rinaldi Rec. Sta.	153
16 July 2007	Chūetsu–Oki	6.6	NIG018	152

Table 4 Lists of the highest vertical vector velocities

Date	*Eq. name*	M_W	*St. name*	V_H (cm s^{-1})
20 September 1999	Chi-Chi	7.6	TCU068	196
20 September 1999	Chi-Chi	7.6	TCU052	138
15 October 1979	Imperial Valley	6.5	El Centro Array #6	80
15 June 2008	Iwate–Miyagi Nairiku	6.9	IWTH25	79
17 January 1994	Northridge	6.7	Rinaldi Rec. Stn.	77

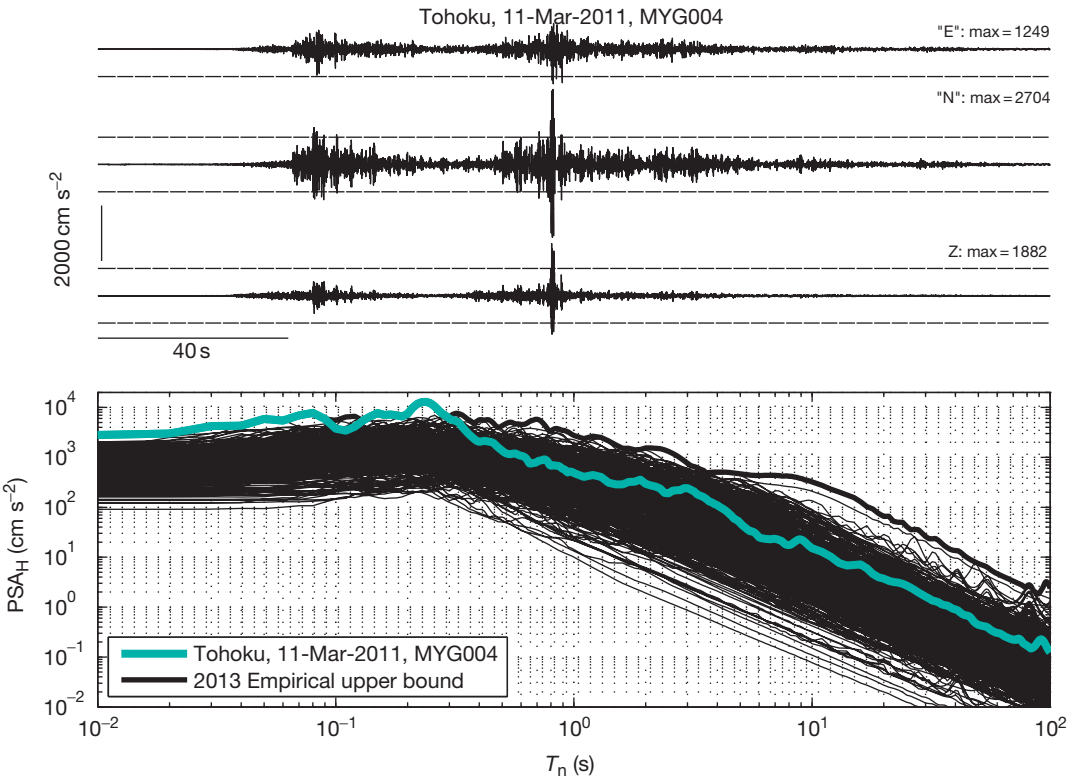


Figure 4 Accelerogram and response spectrum for the MYG004 accelerogram from the 2011 Tōhoku earthquake (**Table 1**). Plotted accelerograms are the central 200 s of the complete ground motion record. In the lower frame, all PSA response spectra (5% damping) from **Figure 3** are reproduced. The spectrum from station MYG004 is superimposed with the light contrasting color.

strong-motion generating areas. **Figure 5** shows the strongest vertical accelerogram listed in **Table 2**. **Figure 6** shows the accelerograms that correspond to the largest horizontal and vertical velocities listed in **Tables 3** and **4**. The response spectrum in **Figure 6** is not one of the highest at short periods, but it forms the empirical upper bound at the longest periods on the figure.

The largest amplitudes of acceleration and velocity observed up to this time will undoubtedly be exceeded on occasions by future events. Is there evidence for substantially more severe motions? Much higher accelerations at very high frequencies have been observed in mines: McGarr et al. (1981) observed 12 g at a distance of 71 m from an earthquake with M_L 0.72 in a deep mine in South Africa. The frequencies associated with this acceleration were between 20 and 400 Hz and so much higher than the frequency range of 1–10 Hz that is more typical of strong motions in earthquake. Associated with surface shaking, there are occasional reports of objects thrown into the air by earthquakes, requiring peak accelerations greater than 1 g; in such cases, the distance thrown sets a lower bound on initial velocity. A famous example is documented by Oldham (1899). Oldham reported that displaced boulders or other stone objects in the 1897 earthquake in Assam, India ($M \sim 8.5$), had been thrown from their original location by distances of up to 2.5 m. Other examples, including eight cases where objects have been thrown from 2 to 4 m, are summarized by Midorikawa (1994). Ohmachi and Midorikawa (1992) suggested that observations of thrown objects in the 1984 western Nagano Prefecture, Japan, earthquake (up to 3.0 m) do not require ground motions substantially greater than the amplitudes represented on **Figures 2** and **3**.

4.20.3 Representation Theorem

The characteristics of strong ground motions that result from slip on a fault are strongly affected by the geometric and dynamic characteristics of the faulting and by the effects of

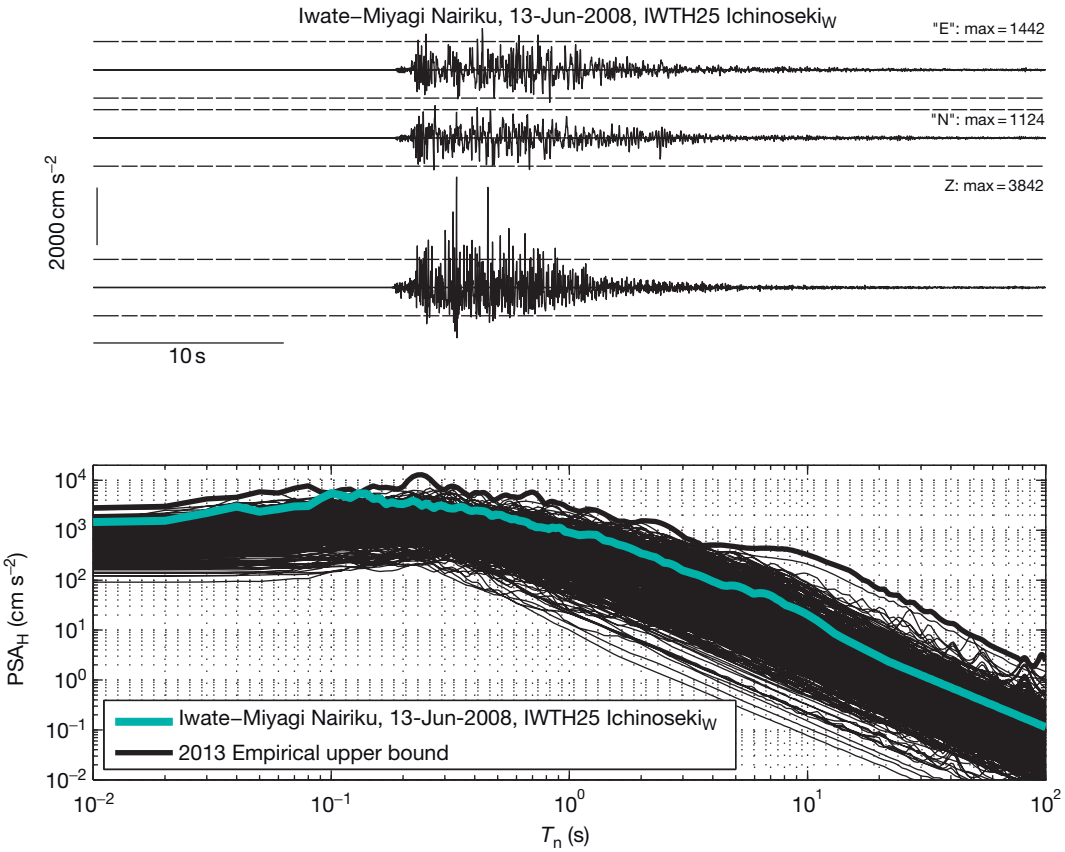


Figure 5 Accelerogram and response spectrum for the IWTH25 (West Ichinoseki) accelerogram from the 2008 Iwate–Miyake Nairiku earthquake (**Table 2**). This accelerogram is the central 50 s of the complete record. The station is located on the hanging wall of the continental thrust fault that caused this earthquake. Other details as in **Figure 4**.

wave propagation through the Earth from the fault to the observation point. The geometric characteristics of faulting include the size, shape, depth, and orientation of the fault area that slipped during the earthquake and the amount, direction, and complexity of slip. The orientation of the fault and the amount and direction of offset may be variable over the fault surface. The important dynamic parameters include where on the fault the rupture initiated (hypocenter), how rapidly it spread over the fault (rupture velocity), how quickly (rise time) and smoothly the slip took place at each point on the fault, and how coherently adjacent points on the fault moved. These dynamic parameters are functions of the stresses acting on the fault, the physical properties of the rock surrounding the fault, and the strength of the fault itself and vary over the surface of a large fault during major earthquakes. Important influences on wave propagation include large-scale geologic structure and layering, lateral variations of structure particularly including basins and shallow geology causing energy to be trapped and/or scattered, and nonlinearity, which can be weak causing only attenuation loss or strong causing significant distortion of the waves.

The ground motion resulting from all of these effects except nonlinearity is expressed mathematically using a representation theorem. Using the notation of Aki and Richards (1980, 2002), one form for the representation theorem is

$$u_n\left(\vec{x},t\right)=\int_{-\infty}^{\infty}\mathrm{d}\tau\iint_{\Sigma}\left[u_i\left(\vec{\xi},\tau\right)\right]c_{ijpq}\nu_j\frac{\partial G_{np}\left(\vec{x},t-\tau,\vec{\xi},0\right)}{\partial\xi_q}\mathrm{d}\Sigma \quad [1]$$

In eqn [1], $u_n\left(\vec{x},t\right)$ gives the nth component of the displacement of the ground at an arbitrary location $\vec{x}$ and at time t. The vector $\vec{\nu}$ is normal to the fault, and the positive direction of the normal defines the positive side of the fault for defining the slip discontinuity. The ith component of the discontinuity in the slip across the fault is given by $\left[u_i\left(\vec{\xi},\tau\right)\right]=u_i^+\left(\vec{\xi},\tau\right)-u_i^-\left(\vec{\xi},\tau\right)$ where $u_i^+\left(\vec{\xi},\tau\right)$ is the displacement on the positive side of the fault, $u_i^-\left(\vec{\xi},\tau\right)$ is the displacement on the negative side of the fault, $\vec{\xi}$ represents a location on the fault surface, and τ is the time that this displacement occurs. The square brackets $[\cdot]$ thus have a special meaning in eqn [1]. The Green's function is given by $G_{np}\left(\vec{x},t-\tau,\vec{\xi},0\right)$. This gives the motion in the n direction at location $\vec{x}$ and time t caused by an impulse point force acting in the p direction at location $\vec{\xi}$ and time τ. Finally, c_{ijpq} gives the elastic constants of the medium. For an isotropic

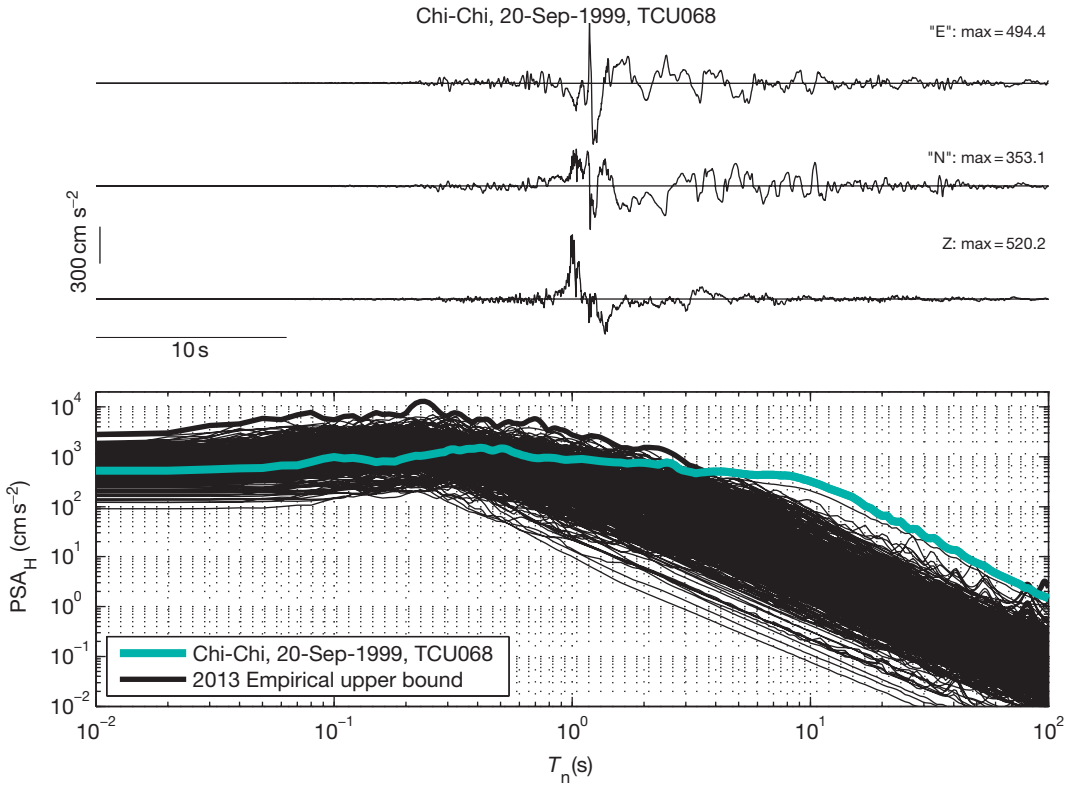


Figure 6 Accelerogram and response spectrum for the TCU068 station recording the 1999 Chi-Chi earthquake (**Tables 3** and **4**). The station is located on the hanging wall, very near the surface rupture, of the continental thrust fault that caused this earthquake. While this accelerogram has been filtered, with special processing, a static displacement can be recovered from this record. This accelerogram is the central 50 s of the complete record. Other details as in **Figure 4**.

medium, $c_{ijpq} = \lambda_E \delta_{ij}\delta_{pq} + \mu(\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk})$, where λ_E and μ are the Lamé constants. By convention, in eqn [1], summation takes place over repeated indices.

Equation [1] represents the ground motion at the site as the linear combination, through the integral over space, of the contributions from each point on the fault surface. The convolution over time incorporates the effect of the rupture at each point taking a finite amount of time to reach its final value. Through the representation theorem, the problem of predicting ground motions requires specification of the offset on the fault as a function of location and time, which incorporates earthquake source physics, and a specification of the Green's function, which incorporates wave propagation.

Equation [1] has certain limitations. For large amplitudes of seismic waves, the stress–strain relationship in the Earth, especially near the surface, becomes nonlinear. In this case, the assumed linear superposition of waves from different parts of the fault does not apply. A common approximation for this case is to predict the motion in 'rock' at some depth beneath the surface, where linearity is assumed to apply, and then treat the wave propagation from that depth to the surface as a vertical propagation problem through the nonlinear medium. On the earthquake source side, eqn [1] assumes that the Green's function is independent of time. This assumption can break down if the faulting process affects the propagation of the seismic waves from the source to the station. For instance, if seismic shear waves are not transmitted through the fault where it is slipping, but they are transmitted through the fault where it is not slipping, then eqn [1] would not hold exactly. In spite of these limitations, eqn [1] forms the mathematical basis for nearly all model-based ground motion predictions and has enabled the inversion of strong ground motion records to obtain models of the slip function for numerous earthquakes.

Isochrone theory is also useful for understanding the relationship of ground motions in the near-source region (where the Green's function is not well approximated by a plane wave leaving the source). The basic theory is given in Bernard and Madariaga (1984) and Spudich and Frazer (1984, 1987). An isochrone is the locus of points on a fault from which the seismic waves all arrive at a selected station at the same time. Thus, in an earthquake, isochrones are different for every station. The isochrone velocity is the velocity at which isochrones move over the surface of the fault. With the implicit assumption that the slip velocity at a point on the fault is discontinuous when the rupture starts and stops, Spudich and Frazer showed that the high-frequency radiation reaching a station is large when the isochrone velocity for waves reaching that station is large or when the slip velocity of the dislocation on the fault is large.

The Fourier transform may be defined as

$$a(\omega) = \int_{-\infty}^{\infty} a(t) \exp(-i\omega t) \mathrm{d}t \qquad [2]$$

where $a(t)$ is a general time series. The real Fourier amplitude spectrum, $|a(\omega)|$, is the magnitude of the complex spectrum. In simplifying the Fourier amplitude spectrum, the phase from the original signal and the ability to recover the original signal through the inverse Fourier transform are lost. The spectrum is usually plotted as a function of the natural frequency, f, rather than the angular frequency $\omega = 2\pi f$.

The Fourier transform of eqn [1] is

$$u_n\left(\vec{x},\omega\right) = \iint_{\Sigma}\left[u_i\left(\vec{\xi},\omega\right)\right]c_{ijpq}\nu_j\frac{\partial G_{np}\left(\vec{x},\omega,\vec{\xi},0\right)}{\partial\xi_q}\mathrm{d}\Sigma \quad [3]$$

Equation [3] is simplified because convolution in the time domain transforms to multiplication in the frequency domain. A common scalar representation of eqn [3] is

$$|u_n(r,f)| = S_F(f)S_{GF}(f) \quad [4]$$

In eqn [4], $S_F(f)$ gives the strength of the excitation of the seismic waves at the fault and $S_{GF}(f)$ describes the effect of wave propagation on the signal once it leaves the source. Both $S_F(f)$ and $S_{GF}(f)$ are Fourier amplitude spectra. The approximation of dropping the spatial integration in going from eqn [3] to [4] is most nearly valid when $|\vec{\xi}| \ll r$, where $r = |\vec{x}|$ is the distance from the origin to the observation point, so that $S_F(f)$ represents the sum of all the energy leaving the fault. Equation [4], and extensions of eqn [4] in which physical processes that affect the Green's function are expressed separately, provides a convenient and widely used way to understand the physical phenomena that influence the observed ground motion.

Based on eqns [1] and [4], the processes controlling strong motion can be understood as the physics of wave propagation from a fault to the site of interest (the Green's function) and the physics of slip on the fault.

4.20.4 Green's Functions

4.20.4.1 Synthetic Green's Functions

Synthetic Green's functions can be computed for any source-station pair, for any focal mechanism, and, conceptually, for any frequency. The limitations are that knowledge of Earth structure limits the maximum frequency that can be meaningfully evaluated and that as the frequency increases, the numerical demands for a realistic Earth structure increase very rapidly.

4.20.4.1.1 Point source in an unbounded medium

The solution for a point source in an unbounded medium is given by simple equations and is useful as background before considering more realistic velocity models. Consider a slip on a small fault represented as a point at the origin. Designate the ground motions at a location $\vec{x}$ as $\vec{u}\left(\vec{x},t\right)$, with $\vec{x} = (r,\theta,\phi)$. The solution uses spherical polar coordinates. The polar angle θ is measured from the normal to the fault plane. The angle ϕ is measured in the fault plane with $\phi = 0$ in the direction of slip on the fault. The vector $\vec{u}$ designates the displacement relative to the static position prior to the earthquake. The motion from an idealized point-source earthquake in a homogeneous medium is expressed as the sum of five terms (Aki and Richards, 1980, 2002):

$$\begin{aligned}\vec{u}\left(\vec{x},t\right) = {}& \vec{u}^{\mathrm{N}}\left(\vec{x},t\right) + \vec{u}^{\mathrm{IP}}\left(\vec{x},t\right) + \vec{u}^{\mathrm{IS}}\left(\vec{x},t\right) \\ & + \vec{u}^{\mathrm{FP}}\left(\vec{x},t\right) + \vec{u}^{\mathrm{FS}}\left(\vec{x},t\right)\end{aligned} \quad [5]$$

The terms in eqn [5] are called, respectively, the near-field term (N), the intermediate-field P-wave (IP), the intermediate-field S-wave (IS), the far-field P-wave (FP), and the far-field S-wave (FS). Equations for these contributions are as follows:

$$\vec{u}^{\mathrm{N}}\left(\vec{x},t\right) = \frac{\vec{A}^{\mathrm{N}}}{4\pi\rho r^4}\int_{r/\alpha}^{r/\beta}\tau M_0(t-\tau)\mathrm{d}\tau \quad [6]$$

$$\vec{u}^{\mathrm{IP}}\left(\vec{x},t\right) = \frac{\vec{A}^{\mathrm{IP}}}{4\pi\rho\alpha^2 r^2}M_0\left(t-\frac{r}{\alpha}\right) \quad [7]$$

$$\vec{u}^{\mathrm{IS}}\left(\vec{x},t\right) = \frac{\vec{A}^{\mathrm{IS}}}{4\pi\rho\beta^2 r^2}M_0\left(t-\frac{r}{\beta}\right) \quad [8]$$

$$\vec{u}^{\mathrm{FP}}\left(\vec{x},t\right) = \frac{\vec{A}^{\mathrm{FP}}}{4\pi\rho\alpha^3 r}\dot{M}_0\left(t-\frac{r}{\alpha}\right) \quad [9]$$

$$\vec{u}^{\mathrm{FS}}\left(\vec{x},t\right) = \frac{\vec{A}^{\mathrm{FS}}}{4\pi\rho\beta^3 r}\dot{M}_0\left(t-\frac{r}{\beta}\right) \quad [10]$$

The radiation pattern terms are the following:

$$\vec{A}^{\mathrm{N}} = 9\sin 2\theta\cos\phi\hat{r} - 6\left(\cos 2\theta\cos\phi\hat{\theta} - \cos\theta\sin\phi\hat{\phi}\right) \quad [11]$$

$$\vec{A}^{\mathrm{IP}} = 4\sin 2\theta\cos\phi\hat{r} - 2\left(\cos 2\theta\cos\phi\hat{\theta} - \cos\theta\sin\phi\hat{\phi}\right) \quad [12]$$

$$\begin{aligned}\vec{A}^{\mathrm{IS}} = {}& -3\sin 2\theta\cos\phi\hat{r} \\ & + 3\left(\cos 2\theta\cos\phi\hat{\theta} - \cos\theta\sin\phi\hat{\phi}\right)\end{aligned} \quad [13]$$

$$\vec{A}^{\mathrm{FS}} = \sin 2\theta\cos\phi\hat{r} \quad [14]$$

$$\vec{A}^{\mathrm{FS}} = \cos 2\theta\cos\phi\hat{\theta} - \cos\theta\sin\phi\hat{\phi} \quad [15]$$

$M_0(t)$ is the cumulative moment released on the fault as a function of time. It is zero before the earthquake begins at $t = 0$. After the fault has stopped slipping, it equals the seismic moment, M_0. The earthquake moment rate, $\dot{M}_0(t)$, is the first time derivative of $M_0(t)$. The radiation patterns for each of the contributions, $\vec{A}$, are vectors, giving both the direction of motion associated with the wave and the wave amplitudes, which depend on the direction from the source.

The moment release on the fault, $M_0(t)$, is a monotonic function, reaching the static moment when slip on the fault ends. Thus, the near-field and intermediate-field terms (eqns [6]–[8]), which are often referred to collectively as 'near-field terms,' carry a static offset. The distance dependence of these terms is r^{-2} or faster. Ground motions associated with the sum of these terms are sometimes referred to as the 'fling step' (e.g., Bolt and Abrahamson, 2003). Because $\alpha \cong \sqrt{3}\beta$, the S-wave term in eqn [8] is likely to dominate in the fling step. Based on

the records that are now available, the verb 'lurch' would better describe these motions. The time derivative, $\dot{M}_0(t)$, returns to zero after the earthquake, so the far-field terms do not contribute to a static offset. The far-field terms (eqns [9] and [10]) have an amplitude dependence of r^{-1}, so at large distances, these terms dominate over the near-field and intermediate-field terms.

4.20.4.1.2 Green's functions for more realistic velocity models

The Green's function has also been solved for a source in a flat-layered medium, where each layer is homogeneous and isotropic (e.g., Luco and Apsel, 1983, presented an approach that minimizes numerical instability). The solution in a layered medium requires numerical evaluation of certain integrals to obtain the Fourier transform of the Green's function (e.g., Apsel and Luco, 1983). These integrals are within the capability of desktop computers.

A realistic Earth model differs significantly from the flat-layered approximation. Most major cities in tectonically active areas are built on sedimentary basins. Basins affect the ground motions, as do topography and seawater (Nakamura et al., 2012; Quinay et al., 2012), and the flat-layered approximation significantly underestimates both the amplitude and the duration of the low-frequency ground motions. The flat-layered approach may be useful as an approximation for the high frequencies in the seismogram. However, for the low frequencies, the preferred approach is to develop a three-dimensional (3D) model for seismic velocities within and outside of the basin and then to solve the equations of motion for an elastic medium directly by a finite-difference or finite-element numerical method (e.g., Bao et al., 1998; Frankel and Vidale, 1992; Furumura and Chen, 2004; Graves, 1996; Hartzell et al., 1999; Hisada and Bielak, 2003; Komatitsch and Tromp, 1999). Komatitsch et al. (2004) showed a good example of how the finite-difference model solution is significantly better than a layer structure at low frequency.

4.20.4.2 Empirical Green's Functions

Hartzell (1978) pointed out that seismograms from a small earthquake solve exactly the wave propagation between the source and the station. Such a seismogram is not strictly a Green's function, since the focal mechanism is included and the rise time of the small event may have some effect, but it is called an empirical Green's function (EGF). The relationship between the seismogram from a small earthquake and a Green's function is explained in the context of eqn [1].

First, consider how to convert the representation theorem (eqn [1]) to a sum of discrete contributions from fault elements. We begin by visualizing that the fault is discretized into K elements ($k=1, 2, \ldots, K$) of equal area, designated as $d\Sigma_k$, as shown in Figure 7.

The dislocation function on the fault can be expressed as a sum of the slip increments. Consider the dislocation at location $\vec{\xi}_k$, which is $u\left(\vec{\xi}_k, \tau\right)$ for an arbitrary time τ. Assume without loss of generality that slip at this point occurs during in the time interval time $0 \le \tau \le T_k$. For numerical application, τ is sampled at the integer multiples of $\delta\tau$ that represents a small time increment. The time from the start of rupture ($t=0$) until it ends may be divided into j_k distinct increments, where $J_\tau = T_k/\delta\tau$. The accumulated slip at time τ ($0 \le \tau \le T_k$) is represented by

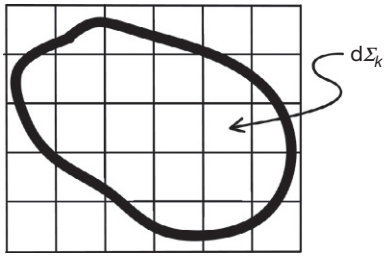


Figure 7 Sketch of a fault rupture, enclosed by the irregular perimeter. This sketch suggests representing the rupture with a regular grid of equal-sized subfaults.

$$\left[u\left(\vec{\xi}_k, \tau\right)\right] = \sum_{j=0}^{n_\tau} \delta u_{jk} \qquad [16]$$

where $n_\tau = \tau/\delta\tau$. The quantity δu_{jk} is an increment of slip that takes place on fault element k in the time interval from $\tau - \delta\tau$ to τ. For a small time interval of duration $\delta\tau$ that takes place at time τ_j and a small fault element of size $\delta\Sigma$ located at $\vec{\xi}_k$, one can write

$$\delta u_{jk} = \left.\frac{\partial u\left(\vec{\xi}_k, \tau\right)}{\partial \tau}\right|_{\tau=\tau_j} \delta\tau \qquad [17]$$

Use of the derivative in eqn [17] is necessary since $u(\xi,\tau)$ gives the cumulative displacement at point ξ, whereas the EGF represents motion for a slip that takes place during a brief time interval on a small fault.

A small earthquake in element k, with moment $M_0^{\mathrm{EGF}_k}$, is to be used as an EGF. While the actual fault dimension of this small earthquake might be estimated from its corner frequency, we define the elemental slip of this event over the area $\delta\Sigma$ as

$$\delta u^{\mathrm{EGF}_k} = \frac{M_0^{\mathrm{EGF}_k}}{\mu\delta\Sigma} \qquad [18]$$

In the representation theorem (eqn [1]), the slip discontinuity is a vector with components $[\delta u_{ki}]$ for $i=1, 2, 3$. The instantaneous magnitude of this vector is

$$|u(\xi,\tau)| = \left(\sum_{i=1}^{3}\left[u_i\left(\vec{\xi}, \tau\right)\right]^2\right)^{1/2} \qquad [19]$$

Let $\vec{\gamma}\ (\xi,\tau)$ be a unit vector in the direction of the slip.

Use of EGFs requires careful accounting for the difference between the sizes of slip elements in eqns [17] and [18]. The approximation used for EGF analysis assumes that the small event is a good approximation for a point source, at location $\vec{\xi}_k$ and at time τ_k, with an impulsive time function. In other words, the slip time function of the EGF is represented by

$$\left[u_i\left(\vec{\xi}, \tau\right)\right] = \left[\delta u^{\mathrm{EGF}_k}\right]\gamma_i\delta(\tau - \tau_k)\delta\left(\vec{\xi} - \vec{\xi}_k\right) \qquad [20]$$

Substituting eqn [20] into eqn [1] implies the notation

$$u_n\left(\vec{x},t\right) = u_n^{\mathrm{EGF}_k}\left(\vec{x},\, t-\tau;\, \vec{\xi}_k\right) \qquad [21]$$

and performing the integration, the result is

$$u_n^{\mathrm{EGF}_k}\left(\vec{x},\, t-\tau;\, \vec{\xi}_k\right) = \left[\delta u^{\mathrm{EGF}_k}\right]\gamma_i\left(\vec{\xi}_k,\tau_k\right)c_{ijpq}\nu_j \frac{\partial G_{np}\left(\vec{x},\, t-\tau_k,\, \vec{\xi}_k,\, 0\right)}{\partial \xi_q}\delta\Sigma \qquad [22]$$

Equation [21] introduces notation that is convenient for subsequent usage and emphasizes that the EGF is the observed seismogram from the small earthquake. Equation [22] shows how the EGF is used mathematically in eqn [1]. Noting that eqn [22] uses the summation convention, one sees how the observed seismogram is represented mathematically by a sum of the mathematical Green's functions. The combination of Green's functions depends on the orientation of the fault ($\vec{\nu}$) and the direction of slip ($\vec{\gamma}$).

By substituting eqn [22] into eqn [1], recognizing that

$$\delta u_{jk} = \frac{\delta u_{jk}}{\delta u^{\mathrm{EGF}_k}}\delta u^{\mathrm{EGF}_k}$$

$\left.\frac{\partial u(\xi_k,\tau)}{\partial \tau}\right|_{\tau=\tau_j}$ can be written as $\frac{\partial u(\xi_k,\tau)}{\partial \tau}\delta(\tau-\tau_k)$, and using the definition in eqn [21], one derives the result that

$$u_n\left(\vec{x},\, t\right) = \sum_{k=1}^{K}\sum_{j=1}^{J_k}\frac{\delta u_k\left(\vec{\xi}_k,\tau_j\right)u_n^{\mathrm{EGF}_k}\left(\vec{x},\, t-\tau_j;\, \vec{\xi}_k\right)}{\delta u^{\mathrm{EGF}_k}} \qquad [23]$$

Multiplying both the numerator and denominator in eqn [7] by $\mu\delta\Sigma$, eqn [7] becomes

$$u_n\left(\vec{x},\, t\right) = \sum_{k=1}^{K}\sum_{j=1}^{J_k}\frac{\left[\dot{M}_{0k}\left(\vec{\xi}_k,\tau_j\right)\mathrm{d}\tau\right]u_n^{\mathrm{EGF}_k}\left(\vec{x},\, t-\tau_j;\, \vec{\xi}_k\right)}{M_0^{\mathrm{EGF}_k}} \qquad [24]$$

where the term in the square brackets is the amount of seismic moment that has been released in element k of the fault in the time interval represented by τ_j during the rupture of the large event. The convolution is represented in eqns [23] and [24] by the sum over j representing the larger rise time on the fault than the rise time of the EGF. This is an essential step to match large earthquake spectra over a broad bandwidth.

The derivation of eqns [23] and [24] is valid with certain assumptions: The first is that the EGF has the same focal mechanism as the earthquake to be modeled. Under this assumption, the normal vector $\vec{\nu}$ and the direction of the dislocation (or rake) vector $\vec{\gamma}$ are the same, so all the terms in the sums in eqns [3] and [22] contribute with the same proportions. The second is that the EGF has a corner frequency higher than any frequency of interest. A third, and generally safer, assumption is that the medium and hence the Green's function do not change with time so that τ_k can be replaced with a general time τ.

There are limitations on the use of the EGF method: The first is that it requires a uniform distribution of small earthquakes, with the desired focal mechanism, to be available over the rupture plane of the hypothetical earthquake to be modeled. The necessary data are rarely available. The second is that it requires the observed small earthquakes to have a good signal-to-noise ratio over the entire bandwidth of the desired synthetic seismogram. Because the spectrum of small earthquakes at low frequencies is very much smaller than the spectrum of large events, many recordings do not have adequate bandwidth. Third, if the ground motion at the station is strong enough to be nonlinear, the linearity assumption behind eqn [1] is violated. Nonetheless, it is a powerful method to keep in mind and use when appropriate.

4.20.4.2.1 Effect of signal-to-noise ratio for EGFs

Figure 8 shows a series of accelerograms for a range of moment magnitudes from 3 to 8 and their corresponding Fourier amplitude spectra. The Fourier spectrum is defined in eqn [2]. At low frequencies, the Fourier amplitude spectrum of acceleration is much smaller at low magnitudes than at high magnitudes. In large magnitude earthquakes, site and basin effects at low frequencies may become major factors in controlling the distribution of damage to significant structures. Thus, predicting the low frequencies well is extremely important. However, Figure 8 shows that the spectral amplitude of a small earthquake at the low frequencies might drop below a useable level, controlled by instrumental or environmental noise, thus preventing that small event from providing useful information in a frequency band of critical interest. Thus, the EGF needs to be selected carefully. Broadband seismographs of the types typically used for high-gain seismic stations may record lower frequencies with a useful signal-to-noise ratio and thus could be installed at sites where EGFs will eventually be needed. Synthetic seismograms generated using eqn [23] or [24] are only valid at frequencies that are well above the noise level in the EGF.

4.20.5 Source Parameters

This section introduces many of the basic parameters that are used to describe the seismic source.

Strong motion cannot be understood without an understanding of the effect of a finite source. In a large earthquake, the fault is tens to hundreds of kilometers long. To evaluate eqn [1], a source model $\left[u_i\left(\vec{\xi},\tau\right)\right]$ that describes the slip on this fault is necessary. Direct observations are impossible, since this slip takes place at depths from the surface to about 15 km or more for shallow crustal earthquakes. For that reason, much of the available information about the space–time behavior of the source comes from inverting eqn [1] for the slip as a function of time and location on the fault. Earthquake source physics is reviewed by Madariaga and Olsen (2002) and in Chapter 4.02 of this comprehensive, and inversions using eqn [1] are discussed in Chapter 4.09 of this Treatise. Therefore, this discussion is brief. Before considering observed results, we describe the Haskell model that captures some of the lowest-frequency characteristics of the slip and establishes essential terminology.

4.20.5.1 The Haskell Model

Haskell (1964) proposed that a propagating ramp is a simple model for slip as a function of time on a fault. The model is

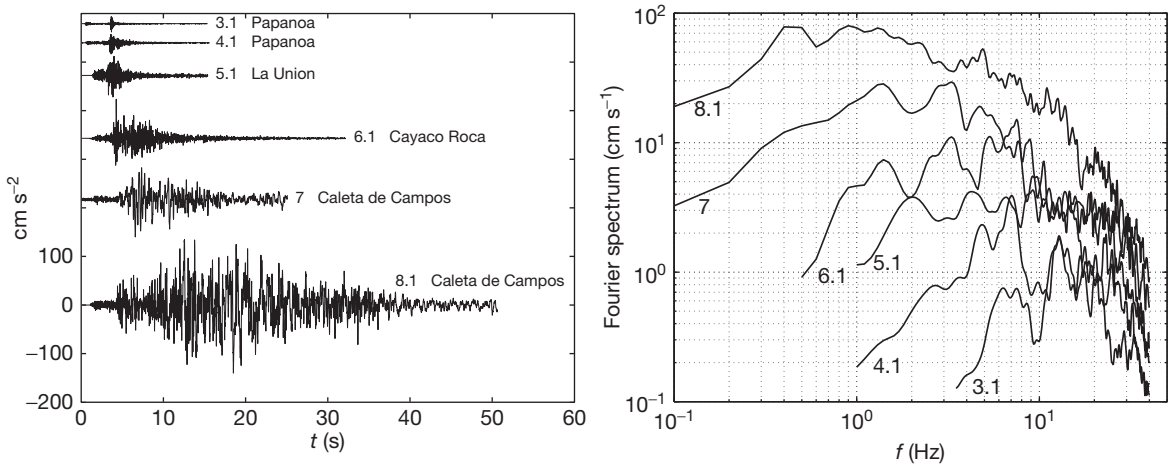


Figure 8 Accelerograms (left) and Fourier amplitude spectra, for a range of event sizes. The earthquakes were chosen to have a hypocentral distance of ~25 km and to sample, as nearly as possible, earthquakes separated by one magnitude unit in size. Modified from Anderson JG and Quaas R (1988) Effect of magnitude on the character of strong ground motion: An example from the Guerrero, Mexico, strong motion network. *Earthquake Spectra* 4: 635–646.

illustrated in Figure 9. The Haskell model is applied to a rectangular fault with length L_E in the x-direction (slip is zero except for $0<x<L_E$) and width W_E in the y-direction (slip is zero except for $0<y<W_E$). On the part of the fault that slips, the time function is

$$[u(x,y,t)] = \begin{cases} 0 & t < \dfrac{x}{v_r} \\ \dfrac{D_E}{t_r}\left(t - \dfrac{x}{v_r}\right) & \dfrac{x}{v_r} \le t \le \dfrac{x}{v_r} + t_r \\ D_E & \dfrac{x}{v_r} + t_r \end{cases} \quad [25]$$

There are five parameters in this model. Fault length (L_E) and fault width (W_E) are defined in the preceding text. The fault slip is D_E. The kinematic part of the rupture is specified by the rupture velocity v_r and rise time t_r. The rupture starts at $x=0$ and progresses across the fault at the rupture velocity v_r. After the rupture front reaches a point x, the fault begins to slip. The slip increases linearly with time to the final offset D_E, and the rise time is the duration during which the slip velocity is non-zero. The normal assumption is that v_r is slightly smaller than the shear velocity, for example, $\sim 0.7\beta$ to $\sim 0.9\beta$.

Equation [25] and Figure 9 describe a rupture that starts at one end of a long fault and propagates in only one direction to the other end. Such a rupture is referred to as unilateral. A bilateral rupture starts in the middle and propagates in both directions simultaneously.

The slip velocity on the fault, $\dot{D}_E = D_E/t_r$, is controlled by the strength of the fault and by the effective stress that is acting across the fault, σ_S. In the Brune (1970, 1971) model,

$$\dot{D}_E = \frac{\sigma_S}{\mu}\beta \quad [26]$$

where β is the velocity of shear waves in the medium. If $\sigma_S = 10$ MPa $= 100$ bars, $\mu = 3.10^4$ MPa, and $\beta = 3000$ m s^{-1}, then $\dot{D}_E = 1$ m s^{-1}. Thus, by this model, exceptional peak velocities of 1–2 m s^{-1} (Figure 2), if not distorted by site effects, suggest that the effective shear stress is on the order of 10–20 MPa for shallow crustal faulting.

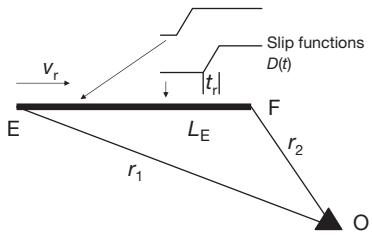


Figure 9 Sketch of a finite fault and definition of parameters to understand the effect of directivity. The heavy line EF represents the fault. Rupture is assumed to start at the epicenter E and propagate toward point F with rupture velocity v_R. The rupture length is L_E. The two slip time functions shown above the fault represent slip at two different points on the fault, as indicated by the arrows. Slip starts and ends sooner for the upper slip function, since it is nearer the epicenter. Slip is not instantaneous, but rather occurs over a time interval of duration t_R. The point O is an arbitrary observation point at distance r_1 from the epicenter and distance r_2 from point F.

4.20.5.2 Example of a Slip Model

Strong-motion data for many strong-motion records have now been inverted to estimate the slip function on the fault as a function of time, using techniques described in Chapter 4.09 of this Treatise. As of 2007, inversions were available for at least 80 earthquakes, in a database of slip functions compiled by Martin Mai (http://www.seismo.ethz.ch/srcmod/Homepage.html). Figure 10 shows the displacement and fault slip velocity for one model (Sekiguchi and Iwata, 2001) of the 1999 Chi-Chi, Taiwan, earthquake (M_W 7.6). This is only one of several models for the Chi-Chi earthquake (e.g., Chi et al., 2001; Ma et al., 2001; Wu et al., 2001; Zeng and Chen, 2001).

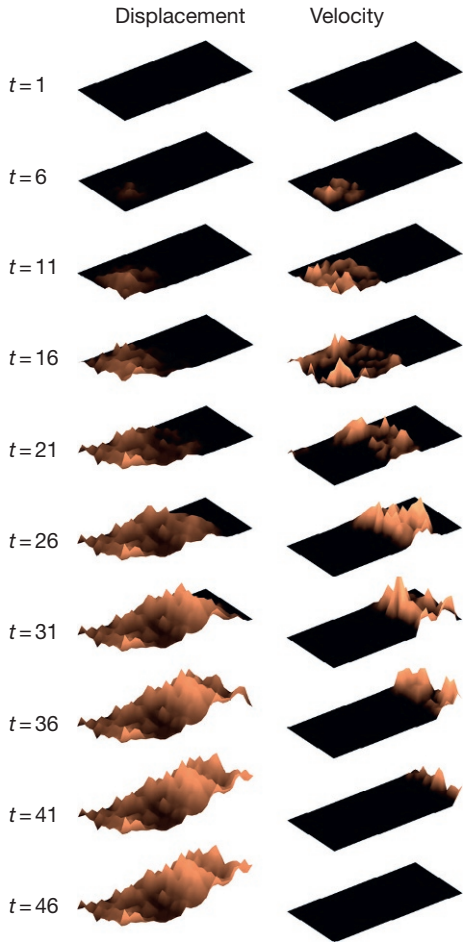


Figure 10 Slip model for the 16 September 1999 Chi-Chi, Taiwan, earthquake (M_W 7.6). This slip model is as determined by Sekiguchi and Iwata (2001) using the strong-motion data from this earthquake. Taiwan is among the best-instrumented countries in the world, and this earthquake is the best-recorded continental thrust earthquake of its size. The fault is 78 km long and 39 km wide. It is oriented approximately north–south, and in this view, the long back edge is near the surface with the north end to the right. The fault dips 300 to the east, and in this view, we are looking at the fault toward the northwest. The column on the left shows total slip on the fault, where the peak (achieved at $t = 31$ s and later) is 9.1 m. The column on the right shows the corresponding velocity. The peak velocity is 2.5 m s^{-1}, in the frame for $t = 31$ s.

These models generally differ in detail, but the broad features are similar. The column on the left, showing displacement, shows that compared with the Haskell model, the final slip is variable over the fault surface. The sequence of inverted slips at different times shows the rupture starting at the hypocenter and spreading initially more or less circularly outward. Later in the event, the slip more closely resembles the propagating ramp model. This feature shows up perhaps more clearly in the velocity column (right). The slip appears somewhat slow to get organized near the hypocenter, but later in the record, it is spatially correlated and behaves more like a propagating ramp moving along the fault, as proposed by Haskell. Unlike the Haskell model, **Figure 10** suggests that the velocity also is rather variable for different points within the slipping portion of the fault. Heaton (1990) summarized several early inversions for fault slip. A key observation is that slip has a relatively short rise time everywhere on the fault, thus confirming that a propagating ramp is a reasonable first-order model.

4.20.5.3 Fourier Amplitude Spectrum of the Seismic Source

The Fourier spectrum of the P- and S-wave radiation from a point source can be obtained by taking a Fourier transform of eqns [14] and [15]. Since the moment function, $M_0(t)$, is a monotonic function, its first derivative is a one-sided pulse. For instance, the Fourier transform of any one-sided pulse is flat at low frequencies and then decreases with increasing frequency above some corner frequency. Using eqn [2], it is easy to show that a boxcar-shaped pulse $p(t) = P_0$ for $0 \le t \le T_p$ and $p(t) = 0$ otherwise has the Fourier amplitude spectrum $p(f) = P_0 T_P(\sin(\pi f T_p)/\pi f T_p)$. At low frequencies, this function is asymptotic to the area under the pulse. It has its first zero at $f_Z = 1/T_P$, and when plotted on logarithmic axes, the high frequency and low-frequency asymptotes intersect at a corner frequency $f_c = 1/(2T_P)$. For any general one-sided pulse, the corner frequency is inversely related to the pulse width. In accord with this general mathematical principle, the Fourier transforms of the far-field P- and S-wave pulses given by eqns [14] and [15] will have low-frequency asymptotes proportional to M_0 and a corner frequency related to the size of the fault.

The model by Brune (1970) is used frequently as a starting point for a model of the source function. By his complete stress drop model, the source term in eqn [4] is

$$S_F(f) = \frac{M_0}{\left(1 + (f/f_c)^2\right)} \qquad [27]$$

where

$$f_c = \frac{2.34\beta}{2\pi a_R} \qquad [28]$$

and a_R is the radius of circular fault. The Brune stress drop is

$$\Delta\sigma_B = \frac{7}{16}\frac{M_0}{a_R^3} \qquad [29]$$

For a suite of earthquakes with a common stress drop, eqns [28] and [29] imply that $M_0 f_c^3 =$ constant. By eqn [27], the Fourier amplitude spectrum of the displacement will fall off with an f^{-2} asymptote at frequencies higher than f_c. Brune (1970) also recognized the possibility of partial stress drop earthquakes, which by his model would have a less rapid falloff, proportional to $\sim f^{-1}$, for some limited range of frequencies at $f > f_c$.

The shape for $S_F(f)$ is important for the development of stochastic synthetic seismograms, as described subsequently in **Section 4.20.8.1.1**. To predict the acceleration spectral amplitudes, eqn [4] is multiplied by $(2\pi f)^2$. That factor may be carried in the source term in the papers describing this approach (e.g., Atkinson, 1993b; Boore, 1983; Chen and Atkinson, 2002). Spectral shapes with multiple slopes are described by Brune (1970, 1976), Savage (1972), Joyner (1984), Atkinson

(1993b), Atkinson and Silva (2000), Chen and Atkinson (2002), and Garcia et al. (2005), among many others.

4.20.5.4 Scaling of Source Parameters

Fault dimensions, seismic moment, moment magnitude, and stress parameters are related, and synthesis of ground motions from a scenario earthquake on a fault, based on eqn [3], should use realistic combinations. The relationships among these parameters are called scaling relations.

Consider an earthquake with rupture length L_E, rupture width W_E, and average slip D_E in rock with shear modulus μ. The seismic moment, M_0, is

$$M_0 = \mu L_E W_E D_E \quad [30]$$

and the moment magnitude, M_W, is

$$M_W = \frac{2}{3}(\log M_0 - \log M_0(0)) \quad [31]$$

Kanamori (1977) introduced the notation in eqn [31] with $M_0(0) = 10^{16.100}$ dyn cm $= 10^{9.100}$ N m.

Based on observed surface rupture and the dimensions of aftershock zones, Wells and Coppersmith (1994) developed an extensive set of empirical relations between fault area, rupture length, slip, and magnitude. The scaling depends on repeat time and tectonic environment (Kanamori and Allen, 1986; Scholz et al., 1986; Wesnousky, 1986), motivating Anderson et al. (1996) to propose equations that used the slip rate and L_E to estimate the magnitude. There are numerous additional studies (e.g., Hanks and Bakun, 2002; Leonard, 2010, 2012; Shaw, 2013). As in **Figure 10**, slip models invariably show high heterogeneous slip, so the definition of length, width, and mean slip can be problematic. Mai and Beroza (2000) defined fault length and width from the autocorrelation distance in the two directions and derived a different set of relations using the results of inversion of the ground motions for 18 earthquakes.

The static stress drop is, from Kanamori and Anderson (1975),

$$\Delta\sigma = b_{SD}\mu\frac{D}{L_C} \quad [32]$$

where L_C is the characteristic length of the smaller rupture dimension. L_C is the fault radius, a_E, for a circular fault or the fault width, W_E, if $W_E \ll L_E$, for a rectangular fault. The constant b_{SD} depends on the fault geometry but is of order unity in all cases. Kanamori and Anderson (1975) gave b_{SD} for three cases: for a buried circular fault $b_{SD} = 7\pi/16$, for a long strike-slip fault that ruptures the surface $b_{SD} = 2/\pi$, and for a long dip-slip fault that ruptures the surface $b_{SD} = (4(\lambda+\mu))/(\pi(\lambda+2\mu))$.

A fault dimension can be estimated from the corner frequency, as described in the previous section, or pulse duration (e.g., Jones and Helmberger, 1998; Kanamori and Rivera, 2004), allowing the determination of the static stress drop by eqn [29]. The approach of Brune (1970) assumes eqn [32] for a circular fault. As a broad generalization, measured stress drop tends to be in the range of 0.1–100 MPa (1–1000 bars) for all sizes of earthquakes (**Figure 11(a)**). **Figure 11(a)**, from Allmann and Shearer (2009), suggests little or no dependence of the static stress drop on magnitude.

Assuming the spectral shape in eqn [27], modified for attenuation at high frequencies as will be discussed subsequently in **Section 4.20.7.14**, McGuire and Hanks (1980), Hanks and McGuire (1981), and Baltay et al. (2013) had measured the stress drop from the flat portion of the high-frequency strong-motion spectrum observed at small distances. The relationship between the spectral level and the RMS stress drop is predicted to be

$$\Delta\sigma_{arms} = a_{rms}\frac{106\rho r}{2A^{FS}(2\pi)^2}\sqrt{\frac{f_c}{f_{max}}} \quad [33]$$

In eqn [33], a_{rms} is the root mean square of the acceleration time history beginning with the S-wave, f_{max} is the observational upper limit of frequencies in the recording, ρ is the density, and r is the source to site distance as in eqn [10]. Since $\Delta\sigma_{arms}$ is determined by the same assumptions as $\Delta\sigma_B$ (eqn [29]), the stress drop estimated from these two methods should correlate closely, and using the excellent data from Japanese strong-motion networks, Baltay et al. (2013) demonstrated that that is the case for r under 20 km, where losses due to attenuation are small.

A different stress parameter that is more related to the dynamics of faulting is the apparent stress, σ_a, defined by Wyss (1970) as

$$\sigma_a = \mu\frac{E_S}{M_0} \quad [34]$$

In eqn [34], E_S is the radiated seismic energy. An equivalent parameter is $\tilde{e} = (E_S/M_0)$ (e.g., Kanamori and Rivera, 2004). Some recent papers that have estimated σ_a include Prejean and Ellsworth (2001), Hwang et al. (2001), Kinoshita and Ohike (2002), Beeler et al. (2003), McGarr and Fletcher (2003), Stork and Ito (2004), Takahashi et al. (2005), and Baltay et al. (2010, 2011). The question of whether or not there is a magnitude dependence of σ_a on M_W has not been settled. Some studies (e.g., Takahashi et al., 2005) have found σ_a increasing with M_W. However, like $\Delta\sigma$, estimates of the apparent stress can be biased toward low values at small magnitudes if attenuation is not corrected perfectly. **Figure 11(b)** shows a summary figure from Baltay et al. (2011), who concluded that there is little or no dependence of σ_a on M_W after special efforts to avoid any contamination by the attenuation of estimates of σ_a. Kanamori and Rivera (2004) showed how $\Delta\sigma_B$ and σ_a are related to each other and to the rupture velocity on the fault.

The 11 March 2011 Tōhoku–Oki earthquake (M_W 9.1) is the best-recorded megathrust earthquake thanks to the extensive networks of strong-motion instruments in Japan. A significant insight from this event is the change in rupture characteristics with depth (Ide et al., 2011; Lay et al., 2012; Ye et al., 2013), illustrated schematically in **Figure 12**. Lay et al. (2012) identified depth-dependent failure domains. Domain A is near the toe of the thrust, generating large slip and tsunamis, but relatively little high-frequency radiation. Domain B has large slip and more short-period energy than domain A. Domain C has patches of slip with high energy radiated at the frequencies of concern in strong motion. These patches are referred to as strong-motion generating areas by Miyake et al. (2003) and dominated the strong-motion records from this earthquake (e.g., Asano and Iwata, 2011; Ide et al., 2011;

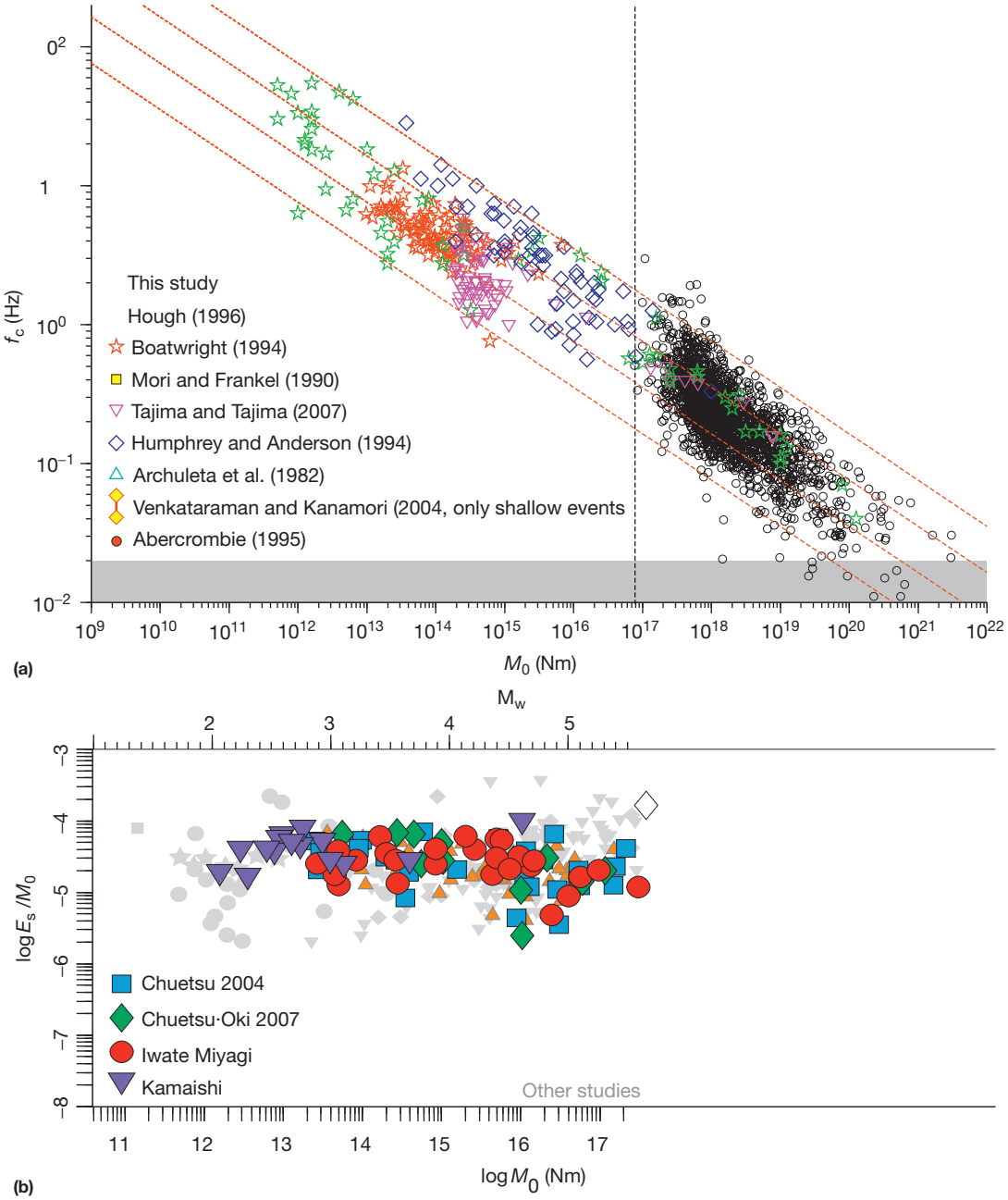


Figure 11 (a) Corner frequency, f_c, and seismic moment, M_0 (reproduced from Allmann BP and Shearer PM (2009) Global variations of stress drop for moderate to large earthquakes. *Journal of Geophysical Research* 114: B01310. http://dx.doi.org/10.1029/2008JB005821). The figure includes results from several local studies, supplemented by a global set of earthquakes studied by Allmann and Shearer (2009). The four diagonal lines show the relationship of M_0 and f_c for four values of static stress drop, $\Delta\sigma$. (b) Ratio of radiated energy, E_S, to seismic moment from a study by Baltay et al. (2011). Independence of E_S/M_0 as a function of M_0 indicates that the apparent stress, σ_a, is also independent of M_0. The figure includes the results of earlier studies by Ide and Beroza (2001) and a study by Baltay et al. (2010).

Kurihashi and Irikura, 2011). Ye et al. (2013) showed how the spectra of smaller earthquakes in different domains differ systematically consistent with this model and that earthquakes associated with failure of the oceanic slab also are highly enriched with high-frequency energy. These phenomena are obvious in the 2011 Tōhoku earthquake because of its enormous size. They are a reminder that average stress parameters like $\Delta\sigma$ and σ_a are not necessarily representative of the behavior of the entire fault. A challenge for future source models is to determine if rupture on smaller, crustal earthquakes also has depth dependence with important implications for strong-motion prediction.

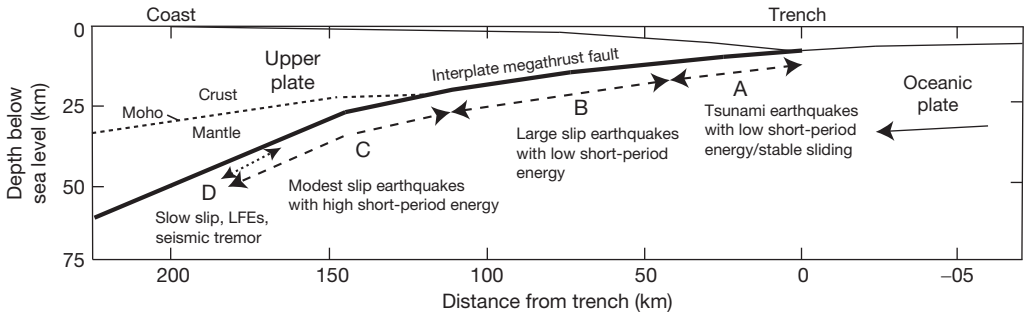


Figure 12 Schematic illustration of the different domains of fault behavior in subduction-thrust zones, motivated by the 2011 Tōhoku earthquake. Modified from Ye L, Lay T, and Kanamori H (2013) Ground shaking and seismic source spectra for large earthquakes around the megathrust fault offshore of northeastern Honshu, Japan. *Bulletin of the Seismological Society of America* 103. http://dx.doi.org/10.1785/0120120115.

4.20.6 The Effects of Fault Finiteness

Large magnitude earthquakes, which generate strong motions, are the result of slip over a large fault area. For instance, based on a regression of available data, Wells and Coppersmith (1994) found the relationship $M_W = 4.07 + 0.98 \log \Sigma_E$, where Σ_E is the area of the fault rupture, given in square kilometers. The purpose of this section is to consider the effect of the finite-fault size somewhat more explicitly than in Section 4.20.5.

4.20.6.1 Spectral Shape

Equations [9] and [10] provide a good starting point for considering the effects of the finite fault, even though the approximation they used in concentrating the total effect of the slip to a point is really only valid at a distance that is large compared to the source dimension. Recalling that rupture will propagate across the fault at a finite speed, typically smaller than the shear velocity, the duration of faulting T_E is controlled by the earthquake dimension, that is,

$$T_E \sim \frac{L_E}{v_r} \qquad [35]$$

From Section 4.20.4.1.1, $\dot{M}_0(t)$ is a one-sided function with a duration T_E. From the theory of the Fourier spectrum, a one-sided pulse transforms to a spectral amplitude that has its peak at zero frequency, is relatively flat until a corner frequency, and then falls off above the corner frequency. The corner frequency is inversely proportional to the pulse duration, that is,

$$f_c \sim \frac{1}{2T_E} \sim \frac{v_r}{2L_E} \qquad [36]$$

A similar consideration leads to the scaling of the corner frequency in the Brune (1970) pulse in eqn [28]. The pulse duration T_E in that model was controlled by the time for the rupture to propagate from the center of the circular fault to the edge and then for the effect of stopping at the edge to propagate back to the center.

Since the area under the pulse is proportional to the moment, eqns [9] and [10] show that as the magnitude of an earthquake increases, the low-frequency spectral level increases. Equation [36] shows that the corner frequency decreases as fault size increases. This scaling law was recognized by Aki (1967). Figure 8 illustrates this effect. To recognize it, though, one needs to recall that Figure 8 shows the Fourier amplitude of acceleration, which is the second derivative of the displacement. The effect of differentiation on the Fourier amplitude spectrum is to multiply by $\omega = 2\pi f$, and double differentiation corresponds to multiplication by ω^2. Thus, where the displacement spectrum is flat, the acceleration spectrum is increasing, and where the displacement spectrum is decreasing as f^{-2}, the acceleration spectrum is flat. Thus, all the spectra in Figure 8 are increasing at the low frequencies and then flattening at the earthquake corner frequency; the corner frequency moves to lower frequencies as magnitude increases. The considerations in this section explain why the magnitude-dependent shapes in Figure 8 are an inevitable consequence of the physics controlling strong motion.

4.20.6.2 Directivity and Duration of Shaking

Directivity is associated with the effects of the finite size of the fault in a large earthquake. A large fault does not rupture instantaneously, as noted in Section 4.20.5.1. Rather, rupture initiates at one point on the fault (the hypocenter) and spreads out from there with a speed that is generally slower than the speed of the S-wave, leading to eqn [35] as an estimate of the duration of faulting. The increase in duration with magnitude is particularly obvious in comparing the accelerograms from the Tōhoku earthquake (Figure 4) with the accelerograms from the Iwate–Miyagi Nairiku and Chi-Chi earthquakes (Figures 5 and 6, respectively). However, the time interval over which direct S-waves are arriving can be either compacted or extended compared to T_E, depending on the combination of the direction of rupture and the location of the station. Directivity impacts both the high- and low-frequency seismograms. At low frequencies, directivity can be responsible for coherent, potentially very destructive pulses with large amplitudes. At high frequencies also, it can cause energy to arrive in short time intervals through its control over the duration of shaking.

To illustrate the effect, consider again Figure 9. Figure 9 envisions an earthquake with rupture length L_E, in which the rupture starts at one end (E, designating the epicenter) and ruptures to the other end at point F. The observation point (O) is located at a distance r_1 from epicenter E and distance r_2 from point F. The rupture is presumed to propagate along the fault at rupture velocity v_r. The time of the first arrival of an S-wave at point O is $t_1 = r_1/\beta$ where β is the shear-wave velocity. The time

of arrival of the last wave at point O is $t_2 = \frac{L_E}{v_r} + \frac{r_2}{\beta} + t_r$. The rise time t_r is expected to be short compared to L_E/v_r. With these assumptions, the duration during which S-wave is arriving at point O, T_s, is

$$T_s = t_2 - t_1 = \frac{L_E}{v_r} + \frac{r_2 - r_1}{\beta} + t_r \qquad [37]$$

For dipping faults, the 3D geometries would lead to a more complicated first-order model for the pulse duration. This analysis makes it clear that for a single site near the fault, where the ground motions are strongest, energy will be arriving from multiple directions, and direct waves from rupture on different parts of the fault will arrive at different times.

Consider in Figure 9 the case where the observation is taken at a point near the origin of the rupture (E), t_r is small, and $v_r = 0.9\beta$. In this case, r_1 is small and r_2 is comparable to L_E, so the time interval over which direct arrivals of S-waves will be arriving is on the order of $2L_E/v_r$. On the other hand, for a station that is located near the far end of the rupture (F), the duration terms in L_E and in r_1 approximately cancel, r_2 is small, and the duration of the seismogram is short. At high frequencies, directivity shows up as a shorter, more intense accelerogram at the far end of the fault, in contrast with a lower-amplitude, longer-duration accelerogram near the origin of rupture. This is illustrated in Figure 13, where the Caleta de Campos record is near the epicenter while the La Union and Zihuatanejo stations are located in the forward directivity direction. Additional examples where directivity has been recognized at high frequencies are given in Frankel et al. (1986), Somerville et al. (1997), and Boatwright and Boore (1982).

Directivity is related to the isochrone velocity (e.g., Spudich and Frazer, 1984). The speed L_E/T_s is the mean isochrone velocity for this site. A high isochrone velocity indicates that energy from a large part of the fault arrives in a short time interval, as occurs near the far end of the fault in Figure 9. Conversely, at a station that is located where the isochrone velocity is low, the seismic energy will be spread over a longer time interval, resulting in lower amplitudes. Consequently, this chapter refers to directions where $(r_2 - r_1)/\beta > 0$ as regions with negative directivity, since the increase in duration tends to lead to a decrease in amplitude, and directions where $(r_2 - r_1)/\beta < 0$ as positive directivity, since conversely, the decrease in duration leads to increased amplitudes.

At low frequencies, directivity associated with propagating rupture causes constructive interference of long-period waves and coherent velocity pulses with large amplitudes. Such pulses have time-function characteristics that are correlated with their polarizations. Near a strike-slip fault, the fault-perpendicular components of ground velocity pulses observed in the region toward which the rupture is propagating are typically two-sided pulses (little net displacement), whereas the fault-parallel components of ground velocity observed near the fault as the rupture passes by are one-sided pulses having significant net displacement associated with the static fault offset, as mentioned in Section 4.20.4.1. This was shown by Archuleta and Hartzell (1981), Luco and Anderson (1983), Anderson and Luco (1983a,b), and Somerville et al. (1997). Somerville et al. (1997) developed empirical ground motion prediction relations that include this effect. Heaton et al. (1995) and Hall et al. (1995) summarized such pulses observed up to that time, which had peak velocities of up to 170 cm s^{-1}, and they showed that such pulses could be especially dangerous to buildings 10–25 stories tall. Figure 14, from Somerville et al. (1997), is selected to illustrate these phenomena. A feature of the intense velocity pulse, as seen in Figure 14, is that it tends to be strongest on the component perpendicular to the fault. This result, which may initially be surprising, is predicted by dislocation modeling. It was first recognized and explained in the interpretation of accelerograms from the 1966 Parkfield, California, earthquake (Aki, 1967; Haskell, 1969).

4.20.7 Effects of Surface Geology

Much attention in strong-motion seismology is devoted to understanding the effects of near-surface geology on the

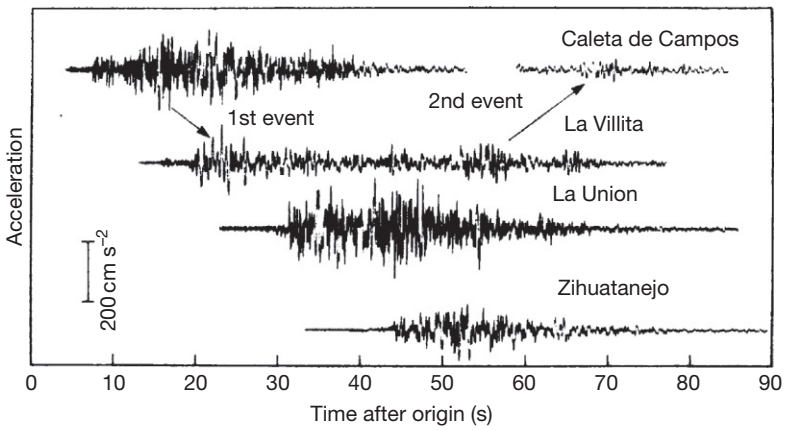


Figure 13 Seismograms from the 19 September 1985 earthquake in Michoacán, Mexico, M_W 8.1. All four stations are above the rupture surface. Caleta de Campos is nearly above the hypocenter of this subduction-thrust earthquake, and the other stations are located on an approximately linear array in the direction of rupture. Caleta de Campos experienced backward directivity. The instrument at Caleta de Campos turned off and retriggered after a 6 s gap as energy from the fault near La Union propagated back to the epicentral area and reintensified the shaking. Note how the duration of shaking decreases at the stations in the direction of forward directivity. Reproduced from Anderson JG, Bodin P, Brune J, Prince J, and Singh SK (1986) Strong ground motion and source mechanism of the Mexico earthquake of September 19, 1985. *Science* 233: 1043–1049.

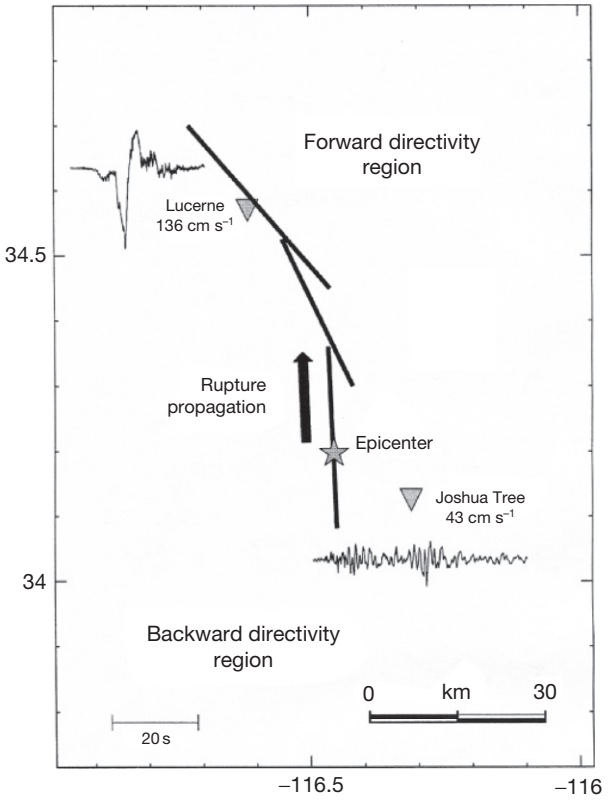


Figure 14 Somewhat schematic illustration of the origin of directivity in the 1992 Landers, California, earthquake (M_W 7.3). The epicenter, identified by the solid star, was near the southern end of the rupture. The strong-motion stations at Joshua Tree and Lucerne, identified by solid triangles, were the two nearest to the rupture. The Lucerne station experienced a coherent velocity pulse with duration under 20 s and peak velocity of 136 cm s^{-1}. The Joshua Tree accelerogram, shown below the station, recorded farther from the fault and in a backward directivity region, had a duration closer to 40 s and relatively low velocities peaking at 43 cm s^{-1}. The timescale for the seismograms is given at the bottom left corner of the figure. Reproduced from Somerville PG, Smith NF, Graves RW, and Abrahamson NA (1997) Modification of empirical strong ground motion attenuation relations to include the amplitude and duration effects of rupture directivity. *Seismological Research Letters* 68(1): 199–222.

ground motions. The near-surface geology is highly variable, so many phenomena are involved and a large number of techniques have been developed for their analysis. This chapter tries to briefly identify the main physical factors affecting strong motion. Kawase (2003), Faccioli and Pessina (2003), and Singh et al. (2003) expanded on the elements of this topic. It is of course understood that in most places, the seismic compressional and shear velocity increase with depth. The 'near surface' is probably best taken to mean the geology above 'seismic bedrock,' where, somewhat arbitrarily, 'seismic bedrock' has shear velocity $\beta \geq 3.0$ km s^{-1} (e.g., Takemura et al., 1995). The main release of seismic energy is within the seismic bedrock.

In uniform rock, joints and microcracks that are open near the surface are closed by increasing pressure with depth. Velocity in laboratory rock samples increases by 10–30% as the pressure is increased from zero to 70–100 MPa (Christensen, 1989; Jaeger and Cook, 1976). Taking a density of 2800 kg m^{-3}, the lithostatic pressure reaches 70–100 MPa at depths of about 2.5–3.5 km. Thus, the influence of the surface extends to at least that depth. Surface outcrops of rock show site effects and thus may not always be suitable for reference ground motions (e.g., Castro et al., 1990; Humphrey and Anderson, 1992; Steidl et al., 1996).

In unconsolidated sediments, compression due to the gravitational load of the overburden will cause both density and the elastic moduli to increase with depth. The shear velocity, $\beta = \sqrt{\mu/\rho}$, also increases because the shear modulus increases more rapidly than the density. Laboratory measurements typically find $\mu \propto (\sigma'_m)^{1/2}$, where σ'_m is the mean principal effective stress (e.g., Kramer, 1996). In the simple case where $\sigma'_m \propto$ depth, the predicted depth dependence of the shear modulus is $\mu \propto (\text{depth})^{1/2}$, and since the density increases relatively slowly with depth, $\beta \propto (\text{depth})^{1/4}$.

Phenomena associated with the propagation of seismic waves between seismic bedrock and the surface include at least all of the following effects: (1) amplification of waves due to the decrease in seismic velocities and density; (2) refraction of body waves toward vertical incidence; (3) reflection and transmission across interfaces; (4) resonance in flat-layered low-velocity layers at the surface; (5) propagation of surface waves; (6) conversion of body waves to surface waves at dipping boundaries, particularly on the edges of basins; (7) trapping and resonance of waves in basins; (8) long linear valleys acting as waveguides; (9) multipathing of body waves causing focusing and high-amplitude ground motions; (10) similar multipathing of surface wave scattering causing an increased duration of ground motions; (11) low coherence of seismic waves (especially higher-frequency body waves) also caused by scattering; (12) 'linear' absorption of energy approximated as low-quality factor Q; (13) scattering by the surface topography; (14) nonlinear stress–strain relationships in near-surface materials; and (15) site–city interactions.

4.20.7.1 NEHRP Site Classifications

The US program in earthquake hazards, called the National Earthquake Hazards Reduction Program (NEHRP), has developed a site classification scheme for the near-surface geology. It is based on the average shear velocity in the upper 30 m, V_{S30}. Setting h_i to be the thickness of a layer and β_i to the corresponding shear velocity, then, in general, the average shear velocity of n layers is

$$\overline{V}_S = \frac{\sum_{i=0}^{n-1} h_i}{\sum \frac{h_i}{\beta_i}} \qquad [38]$$

where the numerator is the thickness of the stack of layers and the denominator is the time for a vertically propagating shear wave to travel through the stack. In eqn [38], the layer starting at zero depth is designated with a zero subscript: thickness h_0 and shear velocity β_0. Then, by definition, $V_{S30} = \overline{V}$ for the special case where $\Sigma h_i = 30$ m. The significance of V_{S30} is that 30 m corresponds to about 100 ft (actually 100 ft = 30.48 m), where 100 ft is a standard depth for geotechnical investigation in the United States, which still uses the English system of units for

common measurements. In a research developed under the US NEHRP, the Building Seismic Safety Council (1997) grouped sites into six 'NEHRP Site Classes' by V_{S30}, and this classification was subsequently adopted as a standard terminology. This classification of near-surface geology is given in Table 5.

4.20.7.2 Amplification Due to the Decrease in Shear Velocity

The site classes in Table 5 are organized by average shear velocity in the upper 30 m. In part, this distinction is based on a theoretical model for how the amplitudes of seismic waves change as a result of propagation into a different kind of material. Setting ρ_0 and β_0 to be the density and shear velocity at the surface and setting ρ_S and β_S to be the density and shear velocity at the earthquake source, assuming density and velocity to vary smoothly and neglecting the free-surface amplification, the predicted gain in amplitude at the surface is given by (e.g., Aki and Richards, 1980, 2002)

$$G_A = \sqrt{\frac{\rho_S \beta_S}{\rho_0 \beta_0}} \qquad [39]$$

Under the NEHRP recommendations for the national building code (Building Seismic Safety Council, 1997), the national zoning map aims to predict ground motion relative to a default site with $v_{S30} = 760\ \text{m s}^{-1}$, that is, on the boundary between NEHRP Classes B and C. The amplification factors, F_a for short-period amplitudes (peak acceleration) and F_v for response spectral amplitude at 1 s period (which is strongly correlated with peak velocity), depend on the predicted response spectral amplitudes S_s and S_1 at those periods, respectively. The amplifications are based at least partly on empirical evidence, such as that developed by Borcherdt and Glassmoyer (1992, 1994). Tables 6 and 7 are included here to show how the code instructs design spectral levels to be modified for different soil classes. Table 6 applies at high frequencies, and Table 7 applies at frequencies centered near 1.0 Hz.

Values of F_a and F_v are plotted in Figure 15. A prominent and interesting feature on Figure 15 is the presumption of strongly nonlinear soil response for the softer soil classes (classes D and E). Essentially, nonlinear response is caused by waves that are too strong to be transmitted through the soil, so they lose amplitude rapidly and permanently deform the soil. Rock and the stiffest soils (classes A, B, and C) are expected to remain linear or very nearly linear. Nonlinear stress–strain relations in soil will be discussed subsequently in this chapter.

The predictions based on eqn [39] are plotted at the zero-amplitude point on the abscissa because they are based on the assumption of linear wave propagation. The relative amplitudes of S_1 in the linear range ($S_1 \sim 0.1$ g) are approximately predicted by eqn [39]. However, the relative amplitudes of S_s in the linear range ($S_s \sim 0.25$ g) are smaller than predicted for site classes C, D, and E. A reasonable explanation for this is that for the high frequencies, attenuation due to low values of Q in the shallow layers could be responsible for the smaller amplifications (see subsequent section). On the other hand, at the intermediate frequencies (1 Hz), attenuation losses will be smaller, so it is expected that the empirical amplifications will better match theory. There is an ongoing research to test, and possibly regionalize, the factors in Tables 6 and 7 (e.g., Sandikkaya et al., 2013).

Table 6 Values of F_a as a function of site classes and mapped short-period maximum considered earthquake spectral acceleration (S_s) in units of g

Site class	$S_s \le 0.25$	$S_s = 0.5$	$S_s = 0.75$	$S_s = 1.0$	$S_s \ge 1.25$
A	0.8	0.8	0.8	0.8	0.8
B	1.0	1.0	1.0	1.0	1.0
C	1.2	1.2	1.1	1.0	1.0
D	1.6	1.4	1.2	1.1	1.0
E	2.5	1.7	1.2	0.9	A
F	[a]	[a]	[a]	[a]	[a]

Note: Use straight line interpolation for intermediate values of S_s.
[a]Site-specific geotechnical investigation and dynamic site response analyses shall be performed.
Source: Building Seismic Safety Council (1997) *NEHRP Recommended Provisions for Seismic Regulations for New Buildings and Other Structures*, 1997 Edition, Part 1: Provisions (FEMA 302), 336 pp. Washington, DC: Building Seismic Safety Council.

Table 5 Definition of NEHRP Site Classes by velocity V_{100} (average shear velocity in the top 100 ft), V_{S30} (average shear velocity in the top 30 m), and blow count

Site class	*Velocity range*	*Blow count range*
A: hard rock	$5000\text{ft s}^{-1} \le V_{100}$ $1500\ \text{m s}^{-1} \le V_{S30}$	Not applicable
B: rock	$2500\ \text{ft s}^{-1} \le V_{100} \le 5000\ \text{ft s}^{-1}$ $760\ \text{m s}^{-1} \le V_{S30} \le 1500\ \text{m s}^{-1}$	$N_{100} \ge 100$
C: very dense soil and soft rock	$1200\ \text{ft s}^{-1} \le V_{100} \le 2500\ \text{ft s}^{-1}$ $360\ \text{m s}^{-1} \le V_{S30} \le 760\ \text{m s}^{-1}$	$50 \le N_{100} < 100$
D: stiff soil	$600\ \text{ft s}^{-1} \le V_{100} \le 1200\ \text{ft s}^{-1}$ $180\ \text{m s}^{-1} \le V_{S30} \le 360\ \text{m s}^{-1}$	$15 \le N_{100} < 50$
E: soil	$V_{100} \le 600\ \text{ft s}^{-1}$ $V_{S30} \le 180\ \text{m s}^{-1}$	$N_{100} < 15$
F: soils requiring site-specific studies	Includes liquefiable soils. Peat or clay with high organic content. High-plasticity clays. Very thick soft/medium stiff clay	

Source: Building Seismic Safety Council (1997) *NEHRP Recommended Provisions for Seismic Regulations for New Buildings and Other Structures*, 1997 Edition, Part 1: Provisions (FEMA 302), 336 pp. Washington, DC: Building Seismic Safety Council.

Table 7 Values of F_v as a function of site classes and mapped short-period maximum considered earthquake spectral acceleration (S_1) in units of g

Site class	$S_1 \leq 0.10$	$S_1 = 0.20$	$S_1 = 0.35$	$S_1 = 0.40$	$S_1 \geq 0.50$
A	0.8	0.8	0.8	0.8	0.8
B	1.0	1.0	1.0	1.0	1.0
C	1.7	1.6	1.5	1.4	1.3
D	2.4	2.0	1.8	1.6	1.5
E	3.5	3.2	2.8	2.4	A
F	[a]	[a]	[a]	[a]	[a]

Note: Use straight line interpolation for intermediate values of S_1.

[a]Site-specific geotechnical investigation and dynamic site response analyses shall be performed.

Source: Building Seismic Safety Council (1997) *NEHRP Recommended Provisions for Seismic Regulations for New Buildings and Other Structures*, 1997 Edition, Part 1: Provisions (FEMA 302), 336 pp. Washington, DC: Building Seismic Safety Council.

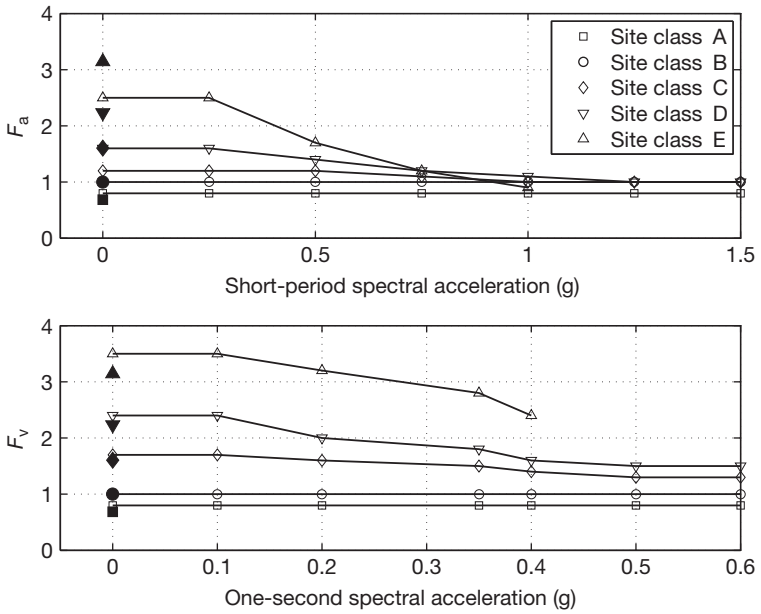


Figure 15 Values of F_a from **Table 6**, and values of F_v from **Table 7**. Solid symbols show the predictions, based on eqn [39], of the relative amplifications of surface ground motions relative to amplifications of site class B. The prediction uses the average of the upper and lower velocity thresholds for each site class range. For site class A, the prediction uses $V_{S30} = 2250$ m s^{-1}, and for site class E, it uses $V_{S30} = 140$ m s^{-1}.

4.20.7.3 Refraction of Body Waves Toward Vertical Incidence

The direction of propagation of seismic waves changes as the wave velocity changes. Snell's law is the fundamental expression that predicts how this will happen. The nature of the change in ray direction on crossing an interface is illustrated in **Figure 8**. The general statement of Snell's law is

$$\frac{\sin j_1}{\beta_1} = \frac{\sin j_2}{\beta_2} \qquad [40]$$

where β_1 is the wave velocity on one side of the interface and β_2 is the wave velocity on the other side of the interface. Angles j_1 and j_2 are measured from the normal to the interface (light vertical line in **Figure 16**).

Figure 17 shows the path that a ray with a large incidence angle will take through a stack of layers with successive properties from engineering bedrock to the site classes A through E. One feature is that in site classes C, D, and E, the energy is turned so that it is propagating very nearly vertical. Because of this feature, it is common in engineering practice to study the effect of surface layers on vertically propagating shear waves. Based on eqns [9] and [10] and confirmed by observations, the S-wave is expected to have a larger amplitude than the P-wave. Since the S-wave particle motion is perpendicular to the direction of travel, a consequence of this near-vertical propagation is that the motion in the S-wave is expected to be predominantly in the horizontal direction on soft sites. A rule of thumb for strong-motion accelerograms is that the peak vertical acceleration is 2/3 the amplitude of the peak horizontal acceleration. In reality, there is a distribution of values, and while it is true that the median ratio is about 2/3, values in the range from 0.5 to 1.0 are all probable. This ratio is perhaps higher than one might expect on soft soils, but the complete seismic wave field is much more complex, as will be discussed subsequently.

4.20.7.4 Reflection and Transmission Across Interfaces

The physics of what happens at an interface is more complicated than what is shown in **Figures 16** and **17**. A detailed

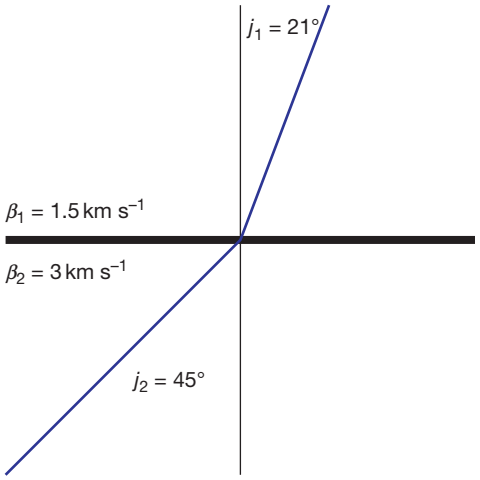


Figure 16 Illustration of change in the direction of a seismic ray on crossing an interface (heavy line). Shear velocities and incident angles are related by Snell's law (eqn [40]).

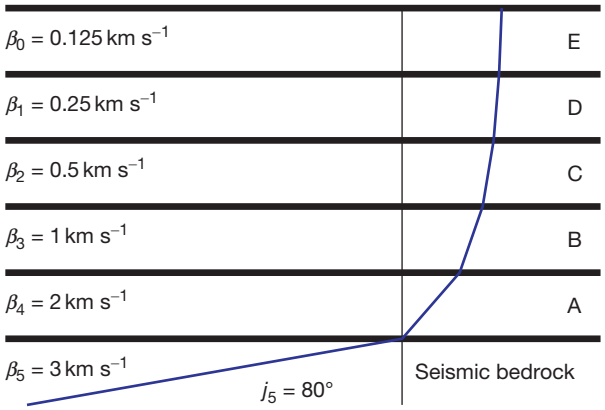


Figure 17 Illustration of the ray path for an S-wave, starting with near-horizontal incidence in seismic bedrock, as the wave is refracted through layers with shear velocities corresponding to the different site classes (**Table 5**) to the surface. Site class is given in the column on the right. Shear velocities (column on the left) are chosen to be representative of the site classes.

review is given by Chapman (2002). For a vertically propagating plane SH-wave incident at the interface, besides the path of the upgoing ray, there is a downgoing reflected ray, as indicated in **Figure 18**. For an SH-wave, assuming Snell's law allows a transmitted wave, the transmission coefficient, which is the ratio of the amplitude of the transmitted wave, A_T, to the amplitude of the incoming wave, A_0, is

$$T_{21} = \frac{A_T}{A_0} = \frac{2\rho_2\beta_2}{\rho_2\beta_2 + \rho_1\beta_1} \qquad [41]$$

and the reflection coefficient, which is the ratio of the amplitude of the reflected wave, A_R, to the amplitude of the incoming wave, is

$$R_{22} = \frac{A_R}{A_0} = \frac{\rho_2\beta_2 - \rho_1\beta_1}{\rho_2\beta_2 + \rho_1\beta_1} \qquad [42]$$

The product $\rho\beta$ is called the impedance. In the limit of $\beta_2 \gg \beta_1$, the transmission coefficient (eqn [41]) approaches

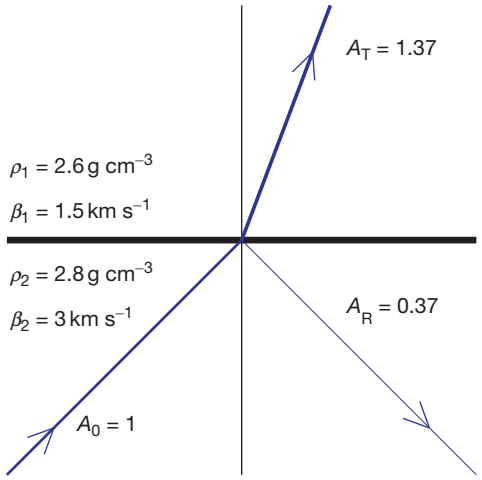


Figure 18 Transmitted and reflected S-waves at an interface. Line weights are proportional to the relative amplitudes of the waves.

doubling the amplitude of the incoming wave. The transmitted energy is not doubled, but rather decreased, because the low-velocity material has a much lower impedance. On the other hand, for transmission from a softer to a harder material, when $\beta_1 \gg \beta_2$, then the transmission coefficient approaches zero and the reflection coefficient (eqn [42]) approaches unity, meaning that the energy is almost totally reflected and is trapped in the low-velocity material. When an SH-wave is reflected from a free surface, the 'free-surface amplification' is a factor of 2. All the energy is reflected, and the wave is described as a sum of the incoming and reflected wave, each with the amplitude of the incoming wave.

4.20.7.5 Amplification and Resonance in Flat, Low-Velocity Layers at the Surface

Multiple layers as in **Figure 17** can induce the amplification of seismic waves much greater than the factor of 2 that limits amplification at a single interface. For instance, if incoming waves encounter two successive interfaces with amplification of, say, 1.37 (as in **Figure 18**), the net effect, which is multiplicative, is a factor of 1.9 amplification. Three interfaces in sequence each with an amplification of 1.37 induce a total amplification of about 2.6. For a smooth decrease in velocity, the amplification was given in eqn [39]. Thus, multiple layers can induce significant amplification, which in turn can cause significant damage (e.g., Hough et al., 1990).

A low-velocity layer near the surface also traps seismic energy and resonance occurs for some frequencies. The phenomenon is usually examined in the context of vertically propagating S-waves, which are a good approximation as seen in **Figure 17**. On entering the low-velocity layer, the S-wave amplitude is increased as given by eqn [41]. The wave is totally reflected at the surface. On encountering the bottom of the low-velocity layer, it is mainly reflected according to eqn [42]. The negative reflection coefficient that arises when the S-wave is reflected back into a low-velocity layer from a layer with higher velocities indicates a 180° phase shift on the reflection.

If the incoming wave is a single-frequency wave, then for certain frequencies, the second peak of the incoming wave is transmitted through the boundary at the right time to be

exactly in phase with the first peak after it has reflected from the surface and the layer boundary. Because of the phase shift in the first peak on the reflection at the base of the layer, the resonance occurs when the wavelength of the wave in the low-velocity layer is four times the layer thickness, that is, when $h_0 = \lambda/4$, where h_0 is the thickness of the top layer and λ is the wavelength. Noting that $\beta = f\lambda$ for a monochromatic shear wave of velocity β and frequency f, the lowest resonant frequency of the layer is

$$f = \frac{\beta}{4h_0} \quad [43]$$

In this model, resonance occurs at a sequence of frequencies, given by

$$f = \frac{(2k-1)\beta}{4h_0} \quad k = 1, 2, \ldots \quad [44]$$

However, because realistic geology generally differs from the mathematical idealization of a perfectly homogeneous, flat-layered medium and because of energy losses, often, the lowest frequency is the only conspicuous resonance in the Fourier amplitude spectrum of a seismogram at the surface.

The response of multiple layers to an incoming plane wave was first solved by Haskell (1953, 1960, 1962). The solution in a linear medium is a fairly standard problem (e.g., Aki and Richards, 1980, 2002) that is easily evaluated by computer. The computer program SHAKE (Schnabel et al., 1972) is one of many widely used in engineering practice that extends the solution to the response of nonlinear soils to vertically incident S-waves. At each layer boundary, high-frequency energy is both transmitted forward and reflected backward. Waves with oblique incidence can be converted from P-waves to SV-waves or from SV-waves to P-waves. Thus, the energy can take many paths to the observer at the surface, resulting in multiple resonance peaks, prolonged duration, and preferential loss of high-frequency energy.

Boore and Joyner (1997) suggested the quarter-wavelength approximation to the response of a layered half-space. This simple approach has been used in engineering applications to find a smoothed average surface amplification function of frequency without any resonant peaks.

The high peak in **Figure 4** appears to be associated with a resonance of this type. When expanded, it consists of four cycles of nearly harmonic motion with the frequency of 4 Hz (Anderson et al., submitted for publication). Consistent with this, earlier and later in the record, one sees predominantly 4 Hz oscillations. Other examples of 1D interpretations of site response include Hough et al. (1990, 1992), Steidl et al. (1996), Hartzell et al. (2000), Satoh et al. (2001a), Gil-Zepeda et al. (2002), Assimaki et al. (2008), and Thompson et al. (2009). Sanchez-Sesma et al. (2011) and Kawase et al. (2011) described a method to invert the autocorrelation of seismic noise at a station for a 1D velocity structure.

4.20.7.6 Trapping Energy in the Form of Surface Waves

Figure 17 should also be considered from the viewpoint of energy in shallow layers that is propagating downward. As depth and shear velocity increase, the ray is bent more toward the horizontal direction, and if the velocity of a layer is too great for the ray to enter while obeying Snell's law, there is total reflection. Thus, seismic energy can be trapped in shallow, relatively low-velocity layers. As the distance increases from the earthquake, this energy organizes into surface waves, which have several features that distinguish themselves from body waves. The theory of surface waves is developed in theoretical seismology textbooks (e.g., Aki and Richards, 1980, 2002).

A few characteristics of surface waves are particularly relevant for strong-motion seismology and are thus worth noting. Because the energy is trapped near the 2D surface of the Earth and propagating in directions parallel to the surface of the Earth, the geometric spreading for surface waves is much less than for body waves that spread over three dimensions. Geometric spreading is proportional to $r^{-0.5}$, in contrast to the body waves for which geometric spreading is proportional to r^{-1} (eqns [9] and [10]). Thus, as distance from the fault increases, surface waves are increasingly dominant in the seismogram. Surface waves that are organized from trapped SH-waves are called Love waves. P- and SV-waves, coupled by conversion from P- to SV- or SV- to P-waves at the boundaries, form Rayleigh waves. The particle motion of a Love wave is horizontal and transverse to the direction of wave propagation. The particle motion for a Rayleigh wave is elliptical (usually retrograde elliptical) and in a vertical plane parallel to the direction of wave propagation. The Love wave motion differs significantly from a vertically propagating S-wave. For the vertically propagating planar S-wave, the motion at the surface is the same everywhere, while a map of Love wave particle motion shows reversals in direction every half-wavelength in the direction of propagation.

Surface wave propagation requires the right combination of frequency and angle of incidence leading to constructive interference of the contributing body waves. The frequency of this constructive interference depends on the incident angle, and thus, surface waves of different frequencies travel with different phase and group velocities. This phenomenon is called surface wave dispersion. Often, low-frequency surface waves travel significantly faster than higher-frequency surface waves. Because of the dispersion and low geometric spreading, surface waves can contribute to damage through their relatively long durations and large amplitudes of ground motions at substantial distances from the source. The dispersion characteristics of surface waves are related to the velocity structure. Because they dominate the background seismic noise, the measurement and interpretation of surface wave dispersion are important tools for estimating the structure of urban basins and strong-motion accelerograph sites (e.g., Horike, 1985; Kawase et al., 1998a,b; Louie, 2001; Malagnini et al., 1997; Matsushima and Okada, 1990; Nazarian and Stokoe, 1984; Sato et al., 1991).

4.20.7.7 Effect of Basin Edges on Incoming Body Waves

Seismic waves incident on a sloping basin boundary will obey Snell's law at the sloping interface. This can result in a wave with a small incidence angle in the rock below the basin being converted to waves with a large incidence angle in the basin (e.g., Frankel, 1993). This is illustrated in **Figure 19**. The first encounter with the sloping basin boundary turns the ray toward the horizontal direction, and after the surface reflection, the second encounter again turns it closer to the

horizontal direction. In this manner, body waves incident on a basin are converted to surface waves at the basin boundary. Figure 19 is two-dimensional but this effect can be three-dimensional. For example, Frankel et al. (2001) found that the strongest motions at some low frequencies (0.125–0.5 Hz) in the Santa Clara Basin are caused by surface waves that are generated at the basin edge and that travel across the basin at an angle almost orthogonal to the direction to the epicenter.

4.20.7.8 Trapping and Resonance in Basins

Understanding the effect of basins on strong ground motions is important because, especially in geologically active regions, cities are generally located in basins. Perhaps the most well-known example of basin response causing severe and damaging ground motions is the Mexico City basin. In the 19 September 1985 Michoacan, Mexico, earthquake (M_W 8.1), the peak acceleration in Mexico City, over 300 km from the fault, was nearly the same as that observed on rock above the fault (Anderson et al., 1986). The ground motion on the SCT record in Mexico City lasted nearly 300 s with a predominant period of about 2 s (Figure 20). Although the peak acceleration is unremarkable on an absolute scale, at a period of 2 s, this is one of the very strongest accelerograms available. Structures in the 10–20 story range were severely stressed by these motions, and many collapsed, resulting in the loss of perhaps 10 000 lives. Another excellent example can be seen in movies showing the progression of strong motion across Japan, as recorded on K-Net and KiK-Net (Furumura et al., 2003, including an electronic supplement). An obvious feature of the movies is the prolonged duration of shaking in the major basins of Japan.

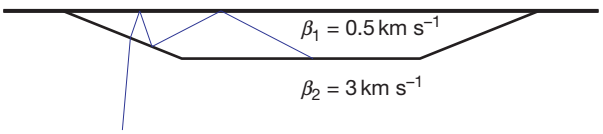


Figure 19 Illustration of the initial path of a ray with nearly vertical incidence, on encountering a low-velocity basin.

The basin effect can be understood as a combination of all the wave-propagation phenomena discussed up to this point. Due to the complexity of the geologic structure, the problem is often investigated using elastic finite-difference, finite-element, or other numerical methods to model the basin response (e.g., Bao et al., 1998; Bard and Bouchon, 1985; Frankel, 1993; Frankel and Vidale, 1992; Graves, 1993, 1996; Olsen et al., 1995a; Pitarka and Irikura, 1996; Pitarka et al., 1994; Sato et al., 1999; Stidham et al., 1999). These methods are typically applied to model frequencies up to about 0.5–1 Hz.

These methods have been applied to understand the low frequencies of strong ground motion. Proceeding first

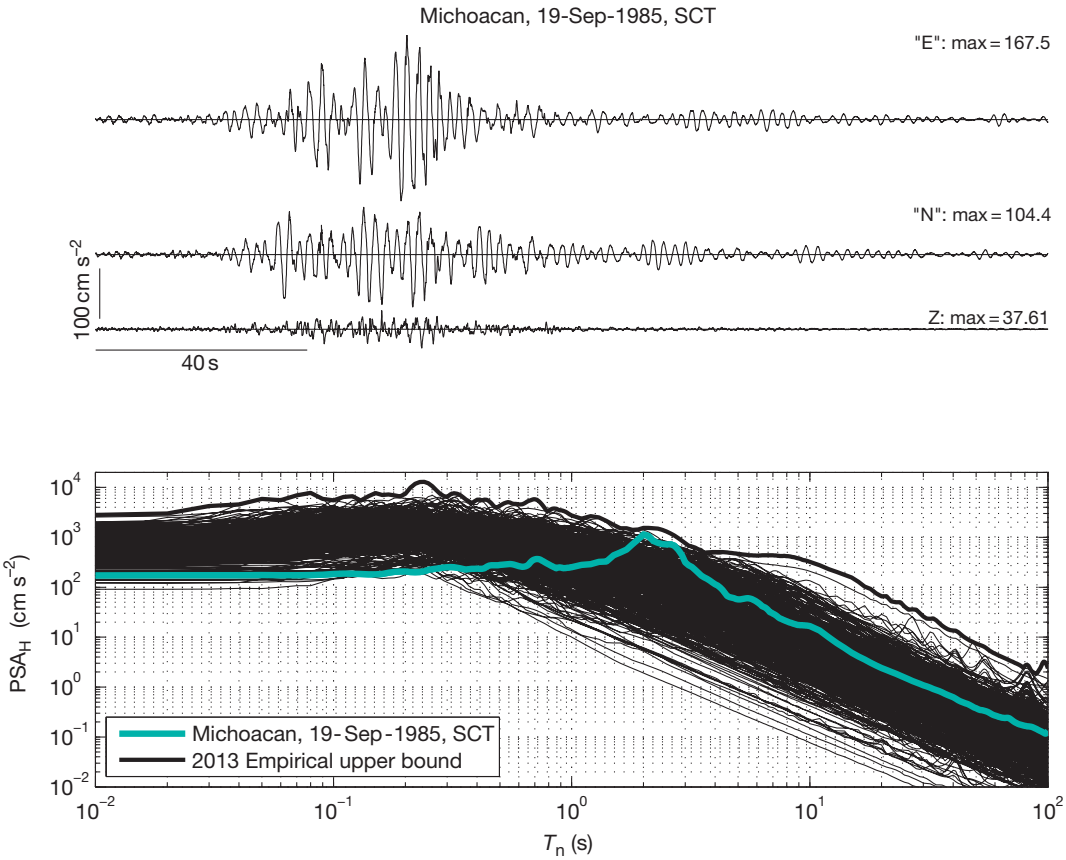


Figure 20 SCT accelerogram recorded in Mexico City, over 300 km from the fault rupture in the 1985 Michoacan, Mexico, earthquake (M_W 8.0). See caption of Figure 4 for more explanation.

clockwise around the Pacific Rim, some of the areas that have been modeled include Wellington (Benites and Olsen, 2005), Kobe (Kawase, 1996; Pitarka et al., 1998), Tottori (Furumura et al., 2003), Tokyo (Sato et al., 1998a,b, 1999), Seattle (Frankel and Stephenson, 2000; Frankel et al., 2002a; Pitarka et al., 2004), Salt Lake City (Olsen et al., 1995b), San Francisco (Graves, 1993), Santa Clara Valley (Frankel and Vidale, 1992; Stidham et al., 1999), Los Angeles (Graves et al., 1998; Hartzell et al., 1999; Olsen, 2000; Olsen and Archuleta, 1996; Olsen et al., 2003; Wald and Graves, 1998), San Bernardino (Frankel, 1993, 1994; Graves and Wald, 2004), Mexico (Furumura and Singh, 2002), Granada (Gil-Zepeda et al., 2002), and Rome (Olsen et al., 2006a,b).

Considering the results of these studies, it is possible to come to some general conclusions about the nature of ground motions in basins: The first is that the 2D and 3D structures are important for trapping energy and increasing the amplitude and duration of the shaking above what is predicted by 1D models (e.g., Hartzell et al., 2000). A second conclusion is that the pattern of high ground motions depends strongly on the direction from which the seismic energy enters the basin and hence the location of the earthquake. A third conclusion is that some basin phenomena remain to be explained, considering, for instance, that the Mexico City observations have not yet been fully modeled. However, perhaps the most important conclusion is that if the basin structure (geometry, P- and S-velocity, density, and Q) is known in sufficient detail and if the earthquake source is also known, then the ground motion in basins can be synthesized at low frequencies. Measuring the basin structure in sufficient detail is not easy, and ground motions are quite sensitive to the velocity model (e.g., Graves and Wald, 2004). Thus, it is essential that basin models be tested and improved. An important challenge is to find efficient ways to use the weak motions from small and distant earthquakes to improve the velocity structure; Iwaki and Iwata (2011) and Tape et al. (2011) had reported progress using two distinctly different approaches.

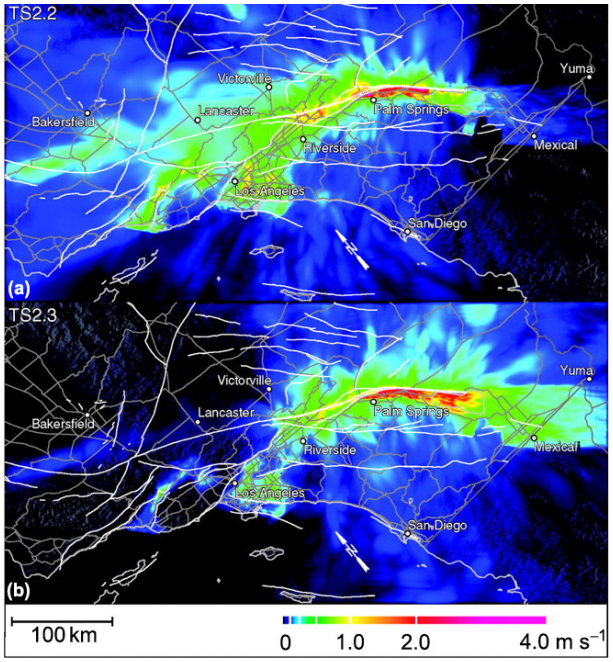


Figure 21 Two maps showing the peak velocity of simulated ground motions for earthquakes on the southern segment of the San Andreas Fault in southern California. Northeast is at the top, and the Pacific Ocean is at the bottom. Other lines are major faults. In both cases, the extent of fault rupture is the same and can be approximately inferred from the portion of the San Andreas Fault highlighted by the large peak velocity in both figures. In case (a), the rupture nucleates at the northwest end (left end) of the rupture and propagates toward the southeast (right). In case (b), the rupture nucleates at the southeast end (right end) of the rupture and propagates toward the northwest (left). The velocity scale on both figures is the same. The color scale at the bottom gives the peak velocity in m s^{-1}. Reproduced from Olsen KB, Day SM, Minster JB, Cui Y, Chourasia A, Okaya D, et al. (2008) TeraShake2: Spontaneous rupture simulations of MW7.7 earthquakes on the southern San Andreas fault. *Bulletin of the Seismological Society of America* 98: 1162–1185.

4.20.7.9 Long Linear Valleys Acting as Waveguides

Numerical simulations have recognized that a long, linear basin can also act as a waveguide, trapping surface wave energy within the basin and guiding it with a relatively small decrease of the amplitude with distance. This phenomenon has not been recognized in actual data, but shows up clearly in synthetic ground motions for an earthquake on the San Andreas Fault (Olsen et al., 2008; **Figure 21**). Olsen simulated ground motions in southern California for two magnitude 7.7 earthquakes on the San Andreas Fault. The final fault slip is the same for both simulations, but in one case, the rupture is simulated to propagate from northwest to southeast and in the other from southeast to northwest. The simulated ground motions for the first case show large amplitudes in the Imperial–Mexicali Valley (**Figure 21(a)**), which is a sediment-filled northward extension of the Gulf of California (Sea of Cortez). Energy is effectively refracted out of the high-velocity mountains on either side of the valley and is guided southeast along the axis of the valley. In the second case, rather than following the San Andreas Fault, the strongest shaking is in the sedimentary basin on the south side of the San Gabriel Mountains (**Figure 21(b)**). Movies showing these and more recent simulations (http://www.scec.org/m8/) make it clear that for the northwest epicenter, energy is trapped and propagated at high amplitude through the southern Imperial Valley and that for the southeast epicenter, energy is trapped and propagated at high amplitude through the essentially continuous sedimentary valley south of the Sierra Madre mountains into Los Angeles.

There are more examples of linear low-velocity waveguides. Li et al. (1990) and Li and Leary (1990) identified trapped waves in low-velocity zones along the San Andreas Fault near Parkfield and in the Oroville fault zone. These waves have subsequently been identified in other fault zones (e.g., see references cited by Ben-Zion, 1998). Also, Shapiro et al. (1998, 2000) found that the accretionary wedge associated with the subduction zone of Mexico efficiently transmits long-period surface waves parallel to the coastline.

4.20.7.10 Multipathing

With complexities in the velocity model, it is obviously possible for some circumstances to exist in which waves take multiple paths to a location, arriving at the same time. When this focusing effect occurs, amplitudes can be unusually large. One case where this has contributed significantly to amplitudes larger than expected is in the 1994 Northridge earthquake, where a lens effect focused energy in the vicinity of Santa Monica (Davis et al., 2000; Gao et al., 1996). Through analysis of data from a dense deployment of instruments to record aftershocks, these studies were able to show that aftershocks with different locations would have energy focused in different locations within Santa Monica in a manner predicted from classical lens theory. Graves et al. (1998) examined the focusing in this region using a finite-element wave-propagation model. A second case is the 1995 Kobe earthquake, where damage was offset from the fault in a location where direct waves arriving through sediments coincided with signals that traveled through rock on the other side of the fault and then converted to surface waves and propagated horizontally to the damage zone (e.g., Pitarka et al., 1997, 1998). In another example, Furumura and Hayakawa (2007) had shown that surface waves from an earthquake north of Tokyo can be focused into central Tokyo.

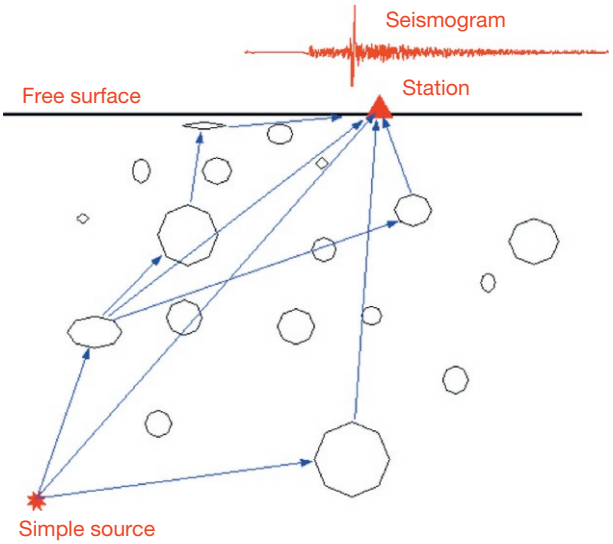


Figure 22 Schematic illustration of how scattering from random velocity irregularities in a medium will generate a coda on a seismogram. Reproduced from Zeng Y (2006) Broadband strong ground motion simulation using a composite source model. In: *Proceedings of the 8th U.S. National Conference on Earthquake Engineering*, San Francisco, CA, 18–22 April 2006, Paper No. 1445.

4.20.7.11 Scattering

Anyone who has spent much time looking at rocks in the field recognizes that geologic formations are not generally homogeneous. Rather, one often sees irregularities in the nature of the rocks on all scales. These irregularities can cause small irregularities in the elastic moduli and in the density, which in turn cause seismic waves to be scattered. The scattering may be weak, with only a small fraction of the energy of a wave being scattered out of the main wave at any one irregularity. However, the net effect is sufficient to cause the coda that appears on the seismograms after the direct waves. The effect is illustrated in **Figure 22**. When each irregularity is excited by the seismic wave, it acts as a weak secondary source radiating waves, strongly preferring S-waves (Aki, 1992; Zeng, 1993), in every direction. Although scattering may be weak, it is sufficient that the amplitudes predicted by the source radiation pattern (eqns [14] and [15]) are not generally observed at high frequencies (e.g., Liu and Helmberger, 1985; Satoh, 2002).

A basic theory for the time dependence of the seismic coda was described by Aki (1969) and expanded by Aki and Chouet (1975), Rautian and Khalturin (1978), and Phillips and Aki (1986). Aki modeled the seismic coda using a single-scattering approximation in a homogeneous medium. Under this model, the amplitude of the coda, $A(f,t)$, is given by

$$A(f,t) = S_F(f) S_{SG}(f) \phi(f,t) \qquad [45]$$

where $S_F(f)$ is the source spectrum as radiated at the fault, $S_{SG}(f)$ is the modification to the spectrum by surface geology (also known as the 'site effect'), and $\phi(f, t)$ depends on how the seismic waves are scattered and attenuated. The functional form for $\phi(f, t)$ is predicted, under the single-scattering model, to be

$$\phi(f,t) = t^{-\nu} \exp\left[-\pi f t / Q_c\right] \qquad [46]$$

where t is the elapsed time after the event, Q_c is coda Q, and $\nu = 1$ for scattering of body waves. The scattering model for the shape of the coda has subsequently been extended to multiple scattering (e.g., Zeng, 1991, 1993; Zeng et al., 1991), scattering in multiple layers (e.g., Yomogida et al., 1997), and multiple scattering including conversion from P- to S-energy (Zeng, 1993). The shape of the coda has also been investigated using an energy diffusion model (e.g., Frankel and Wennerberg, 1987) and by numerical modeling (e.g., Frankel and Clayton, 1984; Menke and Chen, 1984; Richards and Menke, 1983). The numerical studies show how scattering causes attenuation of the amplitude of the direct waves with some energy arriving later, perhaps especially at high frequencies. It should be noted that scattering from major geologic structures contributes very significant energy to the seismic coda (e.g., Spudich and Iida, 1993).

Study of the seismic coda can be useful for strong-motion seismology. According to eqn [46], the time dependence of the coda is practically the same at all of the stations, because at large times, the waves contributing to the coda have all traveled large distances compared to the station separation. Rautian and Khalturin (1978) noted that $\phi(f, t)$ describes the time dependence of the coda for an impulsive source, so with multiple events recorded at a single station, the relative source spectra, $S_F(f)$, can be recovered. Motivated by Tsujiura (1978) and Tucker and King (1984), Phillips and Aki (1986) suggested that records from a single earthquake at multiple stations could be used to estimate the relative site response. Some subsequent studies have supported that conclusion (e.g., Su et al., 1998). Expanding on this idea, Chin and Aki (1991) used differences in the site response between weak and strong motions to identify nonlinearity in site response during the strong motions.

4.20.7.12 Correlation and Coherency

The effects of inhomogeneous geology on seismic waves, including scattering, focusing and defocusing, and travel-time delays, cause ground motions to differ significantly even at stations with relatively small spatial separations. To quantify this, it is useful to bring in the concept of the correlation and coherency of seismograms from nearby stations. In general, when two stations are located near to each other, the seismograms bear a closer resemblance at low frequencies than at high frequencies. Suppose that $u_A(x_A,t)$ and $u_B(x_B,t)$ are two nearby seismograms. The correlation coefficient, C_0, of these two seismograms, over the time interval $t_1 \le t \le t_2$, with no time lag, is estimated by

$$C_0 = \frac{\int_{t_1}^{t_2} u_A(x_A,t)u_B(x_B,t)\mathrm{d}t}{\left\{\int_{t_1}^{t_2} u_A^2(x_A,t)\mathrm{d}t \int_{t_1}^{t_2} u_B^2(x_B,t)\mathrm{d}t\right\}^{1/2}} \quad [47]$$

The time interval can be taken to cover the entire seismogram, just the P-wave, just the S-wave, or other windows of interest.

The correlation coefficient in eqn [47] may be expected to depend on the direction to the source, as waves coming in perpendicular to the line between the two stations will arrive at nearly the same time and thus could have a high correlation. In contrast, waves at other angles will have a phase delay due to the wave-propagation speed that decreases the correlation. A second way to quantify the similarity, especially appropriate for the P- and S-wave windows, is with the lagged correlation coefficient, C_L, which introduces a time shift to the seismograms accounting for this phase delay. Let us suppose that the direct arrival of a certain phase (e.g., the P- or S-wave) at the stations is t_A and t_B, respectively. Then,

$$C_L(t_A - t_B) = \frac{\int_{t_1}^{t_2} u_A(x_A, t-t_A)u_B(x_B, t-t_B)\mathrm{d}t}{\left\{\int_{t_1}^{t_2} u_A^2(x_A, t-t_A)\mathrm{d}t \int_{t_1}^{t_2} u_B^2(x_B, t-t_B)\mathrm{d}t\right\}^{1/2}} \quad [48]$$

When eqn [48] is applied, the resulting correlation coefficient is expected to be less dependent on the direction of wave arrival, and thus, the function that is derived may have more capability to predict the correlation coefficient of these wave packets when waves come from a direction that is not previously observed. If $t_B = 0$ and t_A is considered a variable, the time of the largest positive peak in $C_T(t_A)$ is used to measure the difference in arrival times of a phase at the two stations. Note that for later arrivals, such as the P- or S-wave coda, the incoming waves tend to be highly scattered and thus will be arriving at the station from many different directions. Thus, the time shift for the direct arrivals does not apply to the later arrivals. Consequently, for these later arrivals, eqn [48] is not expected to show much if any greater correlation than what is found in eqn [47].

The coherency is a frequency domain parameter that measures the similarity of seismograms at two nearby stations over a frequency band of width Δf. The separation between the stations is $\Delta x = |x_B - x_A|$, where there is an implicit assumption of spatial stationarity in the region where the measurements are made. After Menke et al. (1990) and Abrahamson et al. (1991), the coherency can be written as

$$\gamma_{AB}(f,\Delta,f,\Delta,x) = \frac{\langle S_{AB}(f)\rangle}{[\langle S_{AA}(f)\rangle\langle S_{BB}(f)\rangle]^{1/2}} \quad [49]$$

The coherence is $|\gamma_{AB}(f,\Delta f,\Delta x)|^2$. For a real time series,

$$S_{AB}(f) = \overline{u}_A(f)u_B(f) \quad [50]$$

is the cross spectrum (the Fourier transform of the cross correlation) of the two seismograms. In eqn [50], $u_A(f)$ and $u_B(f)$ are the Fourier transforms of $u_A(x_A,f)$ and $u_B(x_B,f)$, respectively, and $\overline{u}_A(f)$ is the complex conjugate of $u_A(f)$. The coherency is in general a complex function. For a single $u_A(x_A,t)$ and $u_B(x_B,t)$ and a single frequency, the magnitude of $\gamma_{AB}(f)$ is equal to unity. The notation $\langle S_{AB}(f)\rangle$ is used to indicate that $S_{AB}(f)$ is averaged over the frequency band of width Δf. Abrahamson et al. (1991) smoothed $S_{AB}(f)$, which amounts to the same process. Thus, $\gamma_{AB}(f)$ is calculated by taking an ensemble average. According to Abrahamson et al. (1991), the real part of $\gamma_{AB}(f)$ then gives the similarity of ground motions as a function of frequency with no adjustment for inclined wave propagation, and $|\gamma_{AB}(f)|$ gives the lagged coherency that does include that adjustment. The frequency dependence of the lagged correlation coefficient can be extracted by applying a band-pass filter to the seismograms before evaluation of equations [47] and [48].

In general, a way to present the results when multiple stations are available is to plot $|\gamma_{AB}(f,\Delta f,\Delta x)|$ as a function of the station separation Δx (e.g., **Figure 23**). **Figure 23** shows an example where lagged coherency is high for 1 Hz waves to separations of 100 m, beyond which distance, it starts to decline to a state of essentially no coherency at a separation of 2000 m. As frequencies increase, the coherency decreases. The coherency is highest when direct P- or S-waves are arriving and decreases throughout the coda of the seismogram (e.g., Der et al., 1984; Menke et al., 1990; Vernon et al., 1991). For a site in the Adirondack Mountains of northern New York State, Menke et al. (1990) suggested that the spatial coherence $|\gamma_{AB}(f,\Delta f,\Delta x)|^2 = \exp(-c_\gamma f \Delta x)$, where $c_\gamma = 0.4$–$0.7\ \mathrm{km^{-1}Hz^{-1}}$. Based on a study of coherence in the San Fernando Valley of southern California, Hough and Field (1996) questioned whether this relationship is universal.

The spatial variability of the amplitude is also important. Steidl et al. (1996) suggested that when the input waves from the basement are known from a borehole up to 20 km away from a site of interest and the geologic structure is known at the site of interest, the amplitudes of the surface motions are highly predictable. Hartzell et al. (1997) characterized the standard deviation of ground motion as a function of station separation on a large scale (0–30 km) for stations in the Los Angeles, California, region that recorded aftershocks of the 1994 Northridge earthquake. On average, amplitudes are more variable for larger separations. Tsuda et al. (2006a) looked at predictability of the spectral amplification site response as a function of station separation and the similarity of V_{S30} on a scale of 0–5 km for the dense network of strong-motion instruments in Yokohama, Japan. At low frequencies, when V_{S30} is

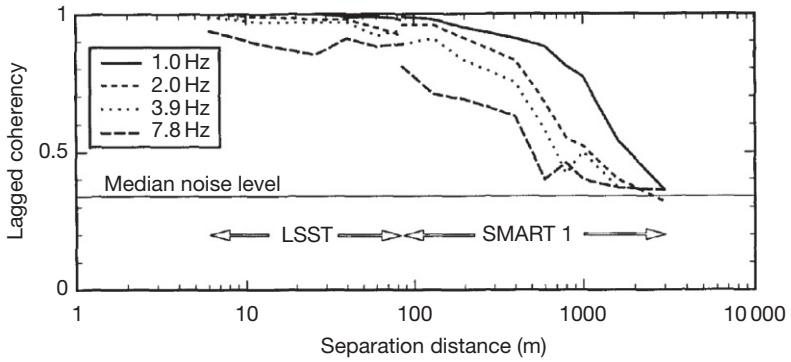


Figure 23 Lagged coherency of ground motions, as a function of station separation, at the SMART 1 array, sited on soft sediments in eastern Taiwan. Reproduced from Abrahamson NA, Schneider JF, and Stepp JC (1991) Empirical spatial coherency functions for application to soil-structure interaction analysis. *Earthquake Spectra* 7: 1–27.

similar, site response there is predictable within a factor of 2 for station separations up to 5 km. The predictability of the spectral site amplification function from nearby stations is lower and less correlated with V_{S30} at frequencies of 5–10 Hz. Additional studies include Goda and Hong (2008), Jayaram and Baker (2009), and Esposito and Iervolino (2011, 2012).

4.20.7.13 Topography

Topography can also have an effect on seismic waves at high frequencies. Consider, for instance, a linear mountain range with a base that is ~10 km wide, bounded by linear valleys on either side. If the shear velocity is 3 km s^{-1}, waves with periods shorter than ~3 s, or frequencies greater than ~0.3 Hz, have wavelengths smaller than the size of the base of the mountain range. As a general rule, when the wavelength is comparable to or smaller than the size of the geologic feature, it is quite reasonable to expect that feature to affect the ground motion. Sánchez-Sesma and Campillo (1991) showed the range of numerical effects of topographic models. A good early observational example is described by Tucker et al. (1984), where the Fourier amplitude spectra at sites on ridges were amplified by as much as a factor of 8 in some frequency bands. In the 2010 Haiti earthquake, Hough et al. (2010) found that along one ridge, the topographic amplification exceeded the nearby sediment-induced amplification. Thus, topography does affect strong ground motions.

The effect of topography on strong motion is, however, not yet well understood, perhaps in part due to the difficulty in numerical modeling of the effects and in part due to the necessity to know complex subsurface geologic structure (Bard and Tucker, 1985). Geli et al. (1988) and Assimaki and Gazetas (2004) provided thorough reviews of research. Geli et al. (1988) found that the amplification of ground motions on mountain tops does occur at wavelengths comparable to the width of the base of the mountains, but that often, the observed amplifications tend to be larger than the predictions for a uniform medium with simple topographic structures on the surface. They attributed the differences to topographic complexity and to complex subsurface geologic structure. Assimaki and Gazetas (2004) and Assimaki et al. (2005) concluded that amplitudes that were too high to be explained by topographic effects alone were related to soil conditions.

Some of the strongest ground motions have been observed on ridges. The Pacoima Dam accelerogram, with peak acceleration of 1.25 g (Trifunac and Hudson, 1971), is in an area of complex topography on a ridge that forms the abutment of a dam. Anooshehpoor and Brune (1989) built a 3D model of foam rubber to study the topographic effects and concluded, consistent with the theoretical models cited by Geli et al. (1988), that the net effect of the topography was small in this case; the frequency of the peak ground motions did not match the frequency of the topographic amplification and the complex topography shielded the site from waves from some parts of the source.

Hartzell et al. (1994) investigated severe damage on Robinwood Ridge in the Loma Prieta, California, earthquake by studying the topographic effects on ground motions in aftershock recordings. They observed very complex effects, including the propagation of significant amounts of energy parallel to the ridge even though the ridge is oblique to the direction from the source. High-frequency spectral peaks can be attributed to local variations in the site response. Although not stated in the paper, the observation of lower-frequency energy traveling parallel to the ridge suggests the possibility that energy is trapped in the ridge, perhaps in some ways analogous to energy trapped in basins, as the ridge also is likely to have lower velocities than the rock at greater depths.

A third important topographic effect was observed in the 1994 Northridge, California, earthquake. The strong-motion station at the Tarzana Cedar Hill Nursery recorded a peak acceleration of 1.78 g in the east–west component. Spudich et al. (1996) installed a dense network of stations to study the regional amplification on aftershock records. They found that the hill does amplify ground motions at 3.2 Hz, about the same frequency as was dominant on the accelerogram, by about a factor of 2 for the east–west component that is parallel to the long axis of the hill (**Figure 24**). The strongest amplifications in the aftershocks, however, were on the north–south motions, transverse to the long axis of the hill.

With the growth in computational capabilities, the models of the effects of topography on ground motions are becoming more sophisticated (e.g., Bouchon and Barker, 1996; Gao et al., 2012; Hayashi et al., 2001; Ohminato and Chouet, 1997). Larsen (2002) included topography in a finite-difference model of ground motions in southern Nevada. This study

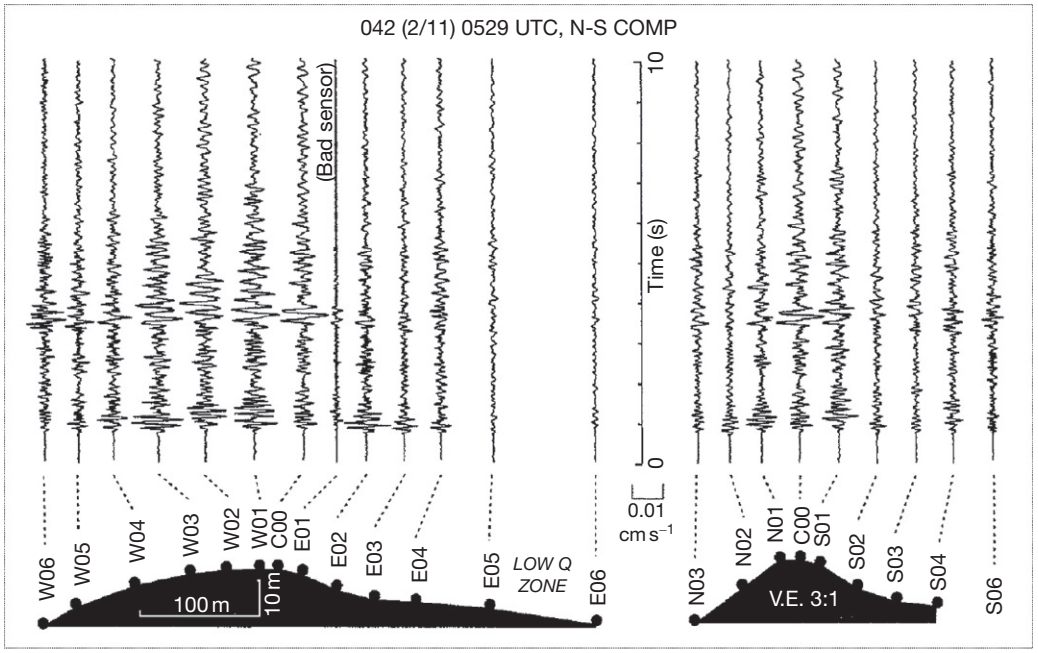


Figure 24 North–south component of seismograms from an aftershock of the 1994 Northridge earthquake, recorded at several stations on a 20 m high topographic feature south of the earthquake. The seismograms on the left show a profile along the long, east–west axis of the hill, and the seismograms on the right show a north–south profile transverse to the hill. Note how the records show higher amplitudes on the hill than at stations at the base of the hill. Reproduced from Spudich PA, Hellwig M, and Lee WHK (1996) Directional topographic response at Tarzana observed in aftershocks of the 1994 Northridge, California, earthquake: Implications for mainshock motions. *Bulletin of the Seismological Society of America* 86(1B): S193–S208.

found that the complex topography of the region (Basin and Range) caused considerable scattering of surface waves. As these capabilities improve, significant progress in the understanding of topographic effects is to be expected.

4.20.7.14 Q

When a seismic wave travels through even a homogeneous medium, it loses a small amount of energy. If the incoming amplitude is A_0, then a plane wave with a single-frequency (f) wave traveling a distance x through the medium with speed β has the amplitude

$$\frac{A}{A_0} = \exp\left(-\frac{\pi x f}{Q\beta}\right) = \exp\left(-\frac{\pi}{Q}\frac{x}{\lambda}\right) \qquad [51]$$

In the distance of one wavelength, λ, the wave energy, which is proportional to the square of the amplitude, decreases by $\exp(-2\pi/Q)$, where the normalization by 2π comes from one complete cycle in radians. Q is thus a dimensionless parameter, where higher Q or quality implies lower losses.

In general, Q is different for P- and S-waves, and it is a function of frequency, lithology, and depth. For the sake of finite-difference modeling at $f<2$ Hz, Olsen et al. (2003) proposed the rule of thumb that, for β in meters per second, $Q_s/\beta=0.02$ for $\beta<1$–2 km s^{-1} and $Q_s/\beta=0.1$ for layers with higher velocities. Using the seismic coda, Aki (1980) found that at frequencies greater than 1 Hz, Q shows clear frequency dependence, which he characterized as

$$Q = Q_0 f^{\delta} \qquad [52]$$

where $0\leq\delta\leq1$ (Figure 25). Similar results were found by Erickson et al. (2004) for Q measured using Lg-waves, which represent S-wave energy trapped in the crust, for the frequency band from 0.5 to 15 Hz. At 1 Hz, typical values of crustal Q for shear waves are in the range of 100–200 in the western United States and about 600–700 in the eastern United States. The high eastern values of Q contribute to the stronger ground motions and larger felt areas of earthquakes in stable regions of eastern North America compared to the tectonically active west, as in the 2011 Mineral, Virginia, earthquake (e.g., Jibson and Harp, 2012). In shallow sediments, shear-wave Q is much lower, in the range of 10–30 or less (e.g., Bonilla et al., 2002; Malagnini, 1996; Seale and Archuleta, 1989).

The spectral decay parameter kappa (κ) has been found useful in engineering studies for characterizing the Fourier spectrum of ground motion. Anderson and Hough (1984) found that the high-frequency Fourier spectrum of acceleration at the surface for frequencies above 2 Hz during earthquakes with $M>5$ behaves as

$$a(f) \sim a_0 \exp(-\pi\kappa f) \qquad [53]$$

as shown in Figure 26. In eqn [53], a_0 gives the intercept of the line but is only relevant to the spectral level for frequencies above the corner frequency. In addition, they found that κ depends on both the station and the distance from the source to the station (Figure 27). They suggested that the parameter κ can be modeled as

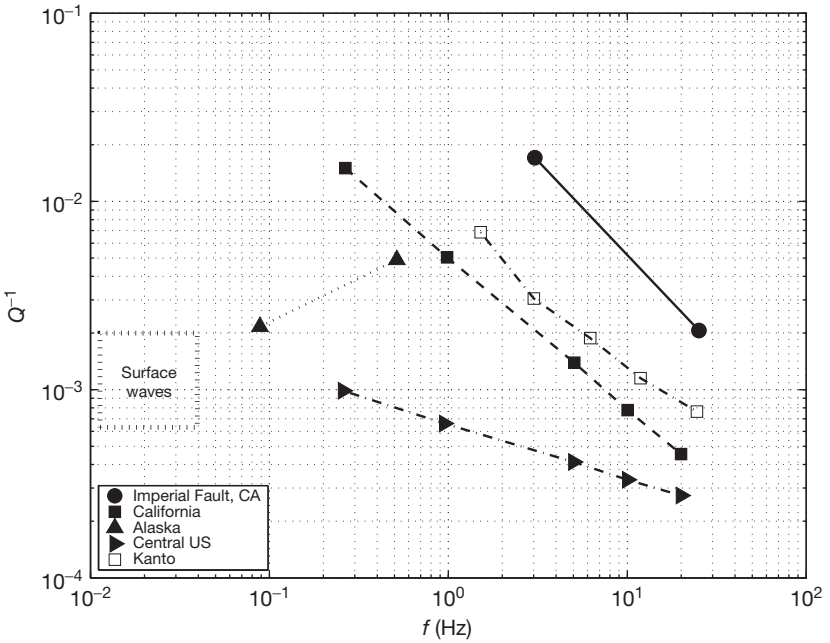


Figure 25 Quality factor Q for seismic coda in several regions of the United States and elsewhere, modified from Cormier (1982), including results of Aki (1980).

$$\kappa = \kappa_0 + \tilde{\kappa}(r_{\text{epi}}) \qquad [54]$$

where κ_0 is a station parameter, r_{epi} is the epicentral distance, and $\tilde{\kappa}(r_{\text{epi}})$ gives a distance dependence related to the entire path. To uniquely partition the observations, they assumed $\tilde{\kappa}(0) = 0$. Subsequently, Purvance and Anderson (2003) found that there does appear to be a source contribution to κ and used the following model:

$$\kappa = \kappa_0 + \kappa_S + \tilde{\kappa}(r_{\text{epi}}) \qquad [55]$$

In eqn [55], κ_S is a source term unique to each earthquake. Again, to uniquely partition the observations, Purvance and Anderson (2003) assumed a zero average of κ_S over all events, that is, $\tilde{\kappa}_S = 0$. They found that the terms for κ_0 and κ_S were both statistically significant with high confidence in explaining the observations in Guerrero, Mexico. Biasi and Anderson (2007) and Kilb et al. (2012) found a similar result in southern Nevada and southern California, respectively.

The parameter κ affects the spectral amplitude in the same way as attenuation loss through eqn [51], that is, κ has the same effect as $x/Q\beta$. Furthermore, the distance dependency implies that it is at least in part related to Q along the entire path. **Figure 28** illustrates a model for how these two separate, and perhaps somewhat independent, contributions to κ might arise. The two rays illustrated in **Figure 28** have low attenuation for propagation in deep, high-Q regions and more attenuation traveling from that high-Q region to the surface. Attenuation traveling to the surface is much stronger in the basin than for a rock site, consistent with **Figure 27**.

If two assumptions were true, κ would be entirely an attenuation parameter: The first is that the acceleration spectrum is flat at high frequencies (i.e., the displacement spectrum falls off as f^{-2} above the corner, as suggested by Aki (1967), Brune (1970), and Hanks (1979, 1982)). This is often referred to as the 'omega-squared model' or ω^{-2} model, using the common symbol for the angular frequency $\omega = 2\pi f$. The second necessary assumption is that Q is independent of frequency. If both were true, then $\kappa = t^*$, where

$$\kappa = t^* = \int_{\text{path}} \frac{\mathrm{d}r}{Q(r)\beta(r)} \qquad [56]$$

The parameter t^* is a standard parameter for introducing Q losses to teleseismic body waves. Unfortunately, as discussed in the preceding text, the two assumptions are not generally valid, but the relationship in eqn [56] is still useful.

The assumption that the average displacement spectrum falls off as ω^{-2} is motivation for the constraint that $\tilde{\kappa}_S = 0$ in modeling data using eqn [55]. Deviation of individual sources from the ω^{-2} model is an explanation for the observations that the κ_S terms in eqn [55] are significant.

As mentioned earlier in the text, Aki (1980) and numerous subsequent papers had found that above 1 Hz, in the crust, Q depends on frequency as in eqn [52]. Anderson (1986) and Hough and Anderson (1988) noted that over limited frequency ranges where eqn [52] is used to model the frequency dependence of Q, and since $0 \le \delta \le 1$, an alternative two-parameter model is

$$\frac{1}{Q_0 f^{\delta}} = \frac{1}{Q_i} + \frac{1}{Q_d f} \qquad [57]$$

where Q_i is a frequency-independent contribution to the total and $Q_d f$ is a frequency-dependent contribution. If eqn [57] is substituted into eqn [53], then, only the term in Q_i affects the shape of the spectrum, which is measured by κ. On this assumption, Hough and Anderson (1988) inverted the distance dependence of κ for a depth dependence of Q_i, with average results shown in **Figure 29**. They suggested that high Q_i occurs at the seismogenic depths in California and that Q_i decreases at

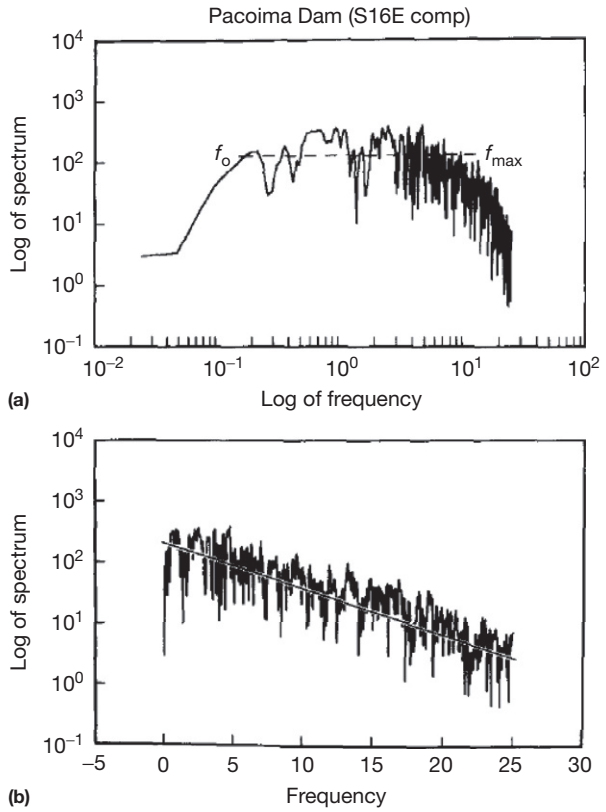


Figure 26 Figure showing measurement of kappa. Fourier amplitude spectrum of the Pacoima Dam accelerogram, 1971 San Fernando earthquake, M_W 6.6. (a) Plotted on typical log-log axes. The corner frequency (f_0 in the notation used for this figure) is identified on this plot where the increasing trend at low frequencies intersects the flat part of the acceleration spectrum, which is indicative of an f^{-2} spectral model. (b) Plotted on semilog axes, on which exponential decay will appear as a straight line (denoted by arrows). The parameter κ was defined by Anderson and Hough (1984) from the slope of this line, fit to the spectrum above the corner frequency. Reproduced from Anderson JG and Hough S (1984) A model for the shape of the Fourier amplitude spectrum of acceleration at high frequencies. *Bulletin of the Seismological Society of America* 74: 1969–1994.

depths where shear strength is lower and creep occurs instead of brittle fracture. Q continues to decrease in the upper mantle, reaching a minimum in the asthenosphere, at depths of 80–220 km (e.g., Durek and Ekstrom, 1996; Lay, 2002).

As κ has become increasingly important for the estimation of strong ground motions, multiple methods to measure it have developed. Anderson and Hough (1984) measured it from the slope of large earthquakes, but large earthquakes are rare and not available for predictions in future earthquakes. For smaller earthquakes, as pointed out by Anderson (1986), the corner frequency moves into the frequency band where κ should be measured. Anderson and Humphrey (1991) and Humphrey and Anderson (1992) proposed to measure it relative to a Brune spectrum (eqn [27]), as a means of separating the effect of the corner frequency and kappa in smaller earthquakes. However, κ and f_0 have a severe trade-off when f_0 is in or near the frequency band where κ is measured, so neither is resolved well and estimates of both are very noisy. For very small events ($M \sim 1$), Biasi and Smith (2001), Biasi and Anderson (2007), Kilb et al. (2012) and Ktenidiou et al. (2014) suggested measuring it from the displacement spectrum, noting that for such small magnitudes, the displacement spectrum is proportional to f^0 in the frequency range of interest (see **Figure 27**), just as the acceleration spectrum is assumed to be for eqn [53]. Comparisons have suggested that these methods are not always equivalent (Kilb et al., 2012).

4.20.7.15 Summary: Spectral Model

Referring to eqn [4], the effect of wave propagation on S-wave spectral amplitudes can be expanded as

$$S_{GF}(r,f) = S_P(r,f)S_{SG}(f)\exp(-\pi\kappa f) \quad [58]$$

In eqn [58], $S_P(r,f)$ gives the effect of the path, which may be interpreted as the propagation through seismic basement rock, from the source to below the site. $S_{SG}(f)$ describes all the effects of propagation through the surface geology from seismic basement to the site, except that it excludes attenuation. Finally, $\exp(-\pi\kappa_0 f)$ gives the effect of attenuation associated with the surface geology at the site, and κ_0 has the same meaning as in eqn [54]. The path term can be further separated as

$$S_P(r,f) = g_{GS}(r)\exp\left(-\frac{\pi r f}{Q(f)\beta}\right) \quad [59]$$

where $Q(f)$ takes on the more frequently used functional form of eqn [52] and $g_{GS}(r)$ describes the geometric spreading of the phase. For instance, for body waves in an infinite homogeneous medium, $g_{GS}(r) = r^{-1}$ (e.g., eqn [10]). Some of the papers that utilize this approach to describe the spectra of ground motions include Margaris and Boore (1998), Malagnini and Herrmann (2000), Malagnini et al. (2000a,b), Akinci et al. (2001), Malagnini et al. (2002), Bay et al. (2003), Malagnini et al. (2004), and Scognamiglio et al. (2005). Additional spectral models are cited in **Section 4.20.8.1.1**.

4.20.7.16 Nonlinearity

The usual assumption in wave propagation is that the amplitudes of strains are small and that the restoring stress in the deformed material is proportional to the strain applied by the wave. However, it is a common experience that soils can be deformed permanently by various straightforward methods (e.g., compacting by stomping with a boot). In the case of compacting soils, the deformation exceeds the elastic limit, and as a consequence, the elastic forces are not sufficient to cause the soil to return to its original shape. Under strong ground motion, it is possible for the waves to exceed the elastic limit of the material, where restoring stress is no longer proportional to the strain. Kramer (1996) provided a good textbook review of the topic. Among the first strong-motion records that convincingly observed nonlinearity, one can cite the 1964 Niigata earthquake, M_W 7.6 (Kawasumi, 1968; Kudo et al., 2000); the 1989 Loma Prieta earthquake, M_W 6.9 (Chin and Aki, 1991); and the 1994 Northridge earthquake, M_W 6.7 (Field et al., 1997). The accelerogram from the 1964 Niigata earthquake is shown in **Figure 30**. About 10 s into this record, the amplitudes of high-frequency waves abruptly drop as a

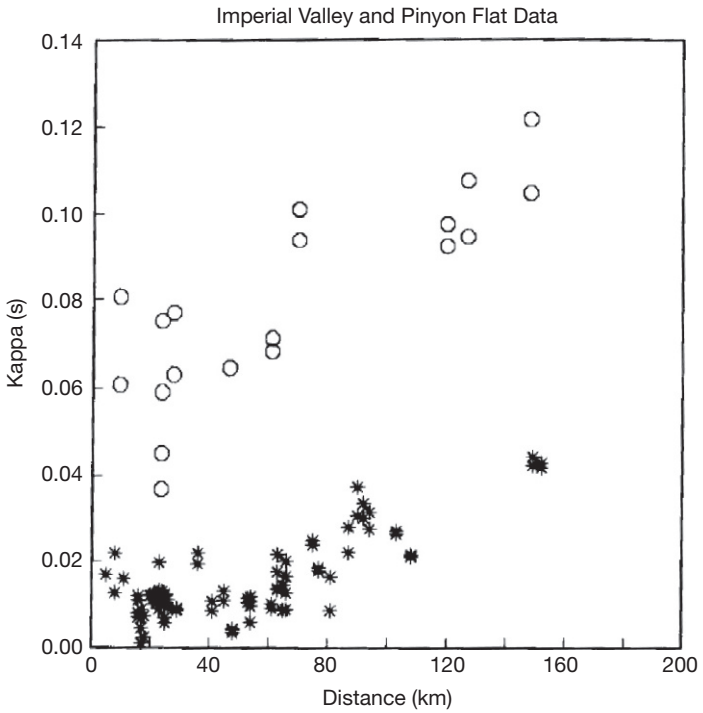


Figure 27 Kappa versus distance at two stations. Both stations are located in southern California. The Imperial Valley station ('o' symbols) is located in El Centro, on deep sediments. The Anza station ('*' symbols) is located 120 km northwest, on granite of the Peninsular Range. The distance dependence is similar, which Anderson interpreted as due to wave propagation affecting seismograms for both stations in the same way. The intercept is shifted, which Anderson interpreted as due to different values of (trial mode) in the shallow geologic materials immediately below the stations (e.g., the upper 5 km). Reproduced from Anderson JG (1986) Implication of attenuation for studies of the earthquake source. In: *Earthquake Source Mechanics*, Vol. 6: *Maurice Ewing*, pp. 311–318. Geophysical Monograph 37. Washington, DC: American Geophysical Union.

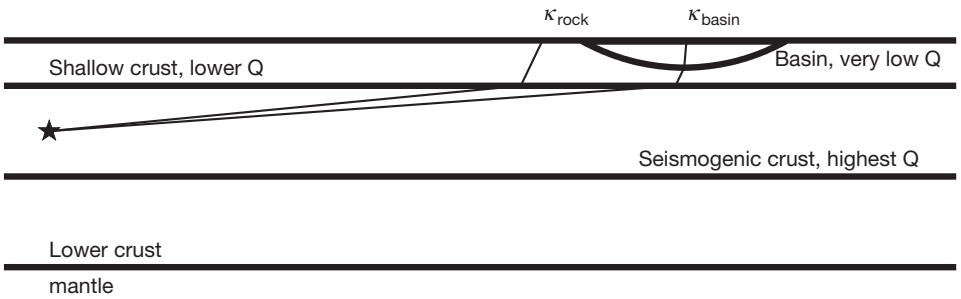


Figure 28 Cartoon illustrating how kappa depends on both distance and site conditions. The figure represents a three- or four-layered crust overlying the mantle. The star represents an earthquake source in the seismogenic crust, with rays emanating from the source to a rock station and a station in a basin. The model for κ in eqn [54] envisions high-Q propagation in the seismogenic and lower crust, resulting in only a small increase in κ for this part of the path. It envisions most of the attenuation along the path taking place in the shallow crust and even higher attenuation taking place for ray paths that travel through a basin instead of only rock in the shallow crust.

result of liquefaction below the station. The continuing long-period energy is considered by Kudo et al. (2000) to be a good example of basin amplification, resulting in the very large spectral amplitudes at periods of 6–8 s.

In the linear models, the only material constants are the elastic moduli. In the nonlinear case, every material can be nonlinear in its own way, as the stress is no longer proportional to the strain. Figure 31 shows one commonly used shape of nonlinear stress–strain curve and defines some important variables. Those are the strain and stress at the corner, σ_C and ε_C, respectively; the low-strain shear modulus, μ_0; the secant shear modulus $\mu_{SEC} = \sigma_C/\varepsilon_C$; and the tangent shear modulus that depends on the local slope of the curve. The ratio of the secant elastic modulus to the low-strain shear modulus, μ_{SEC}/μ_0, is called the modulus reduction. Figure 32 shows the secant modulus as a function of shear strain, corresponding to the stress–strain curve in Figure 31. The area inside the hysteresis loop of Figure 31 is proportional to the energy loss in one deformation cycle. The damping ratio is defined by

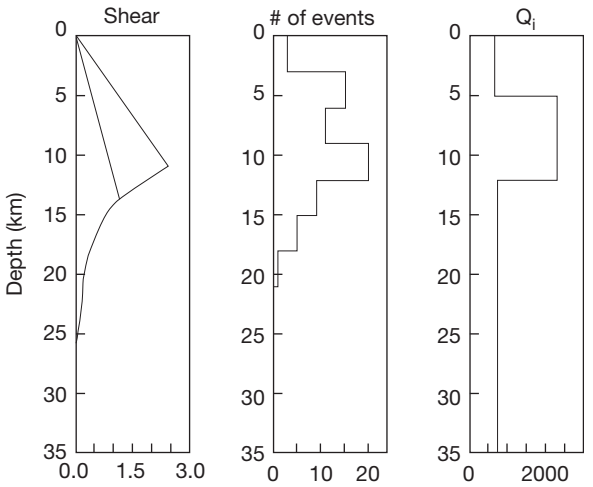


Figure 29 Theoretical maximum shear resistance as a function of depth (after Sibson, 1984), depth of earthquakes near Anza, California (after Doser and Kanamori, 1986), and schematic estimate of Q_i for an idealized three-layer model proposed by Hough and Anderson (1988). Shear resistance is in kilobars. Reproduced from Hough SE and Anderson JG (1988) High-frequency spectra observed at Anza, California: Implications for Q structure. *Bulletin of the Seismological Society of America* 78: 692–707.

$$\zeta = \frac{W_D}{4\pi W_S} = \frac{1}{2\pi}\frac{A_{loop}}{\mu_{SEC}\varepsilon_C^2} \quad [60]$$

where W_D is the dissipated energy, W_S is the maximum strain energy, A_{loop} is the area inside one complete cycle of the hysteresis loop, and ε_C is the peak strain in the deformation cycle (Kramer, 1996). Comparing with the definition of Q (eqn [51]; see also Kramer, 1996), the damping ratio is related to Q by

$$Q = \frac{1}{2\zeta} \quad [61]$$

Figure 32 thus also shows the quality factor Q corresponding to this model. In this model, Q is independent of frequency. The measurements of Q in soils cited in Section 4.20.7.16 were taken at small amplitudes and are much smaller than predicted by this model, perhaps indicating that additional phenomena contribute to the overall attenuation in soils besides the nonlinearity in the typical engineering models.

Through the analysis of a vertical array of accelerometers, it is possible to infer the stress–strain relationships of the materials between pairs of instruments (e.g., Pavlenko and Irikura, 2002; Pavlenko and Loh, 2005). Figure 33 shows an example for the site at Port Island that experienced strong liquefaction during the 1995 Kobe, Japan, earthquake. One important feature is that nonlinear soil response can follow a different pattern from that shown in Figure 31. In particular, at early

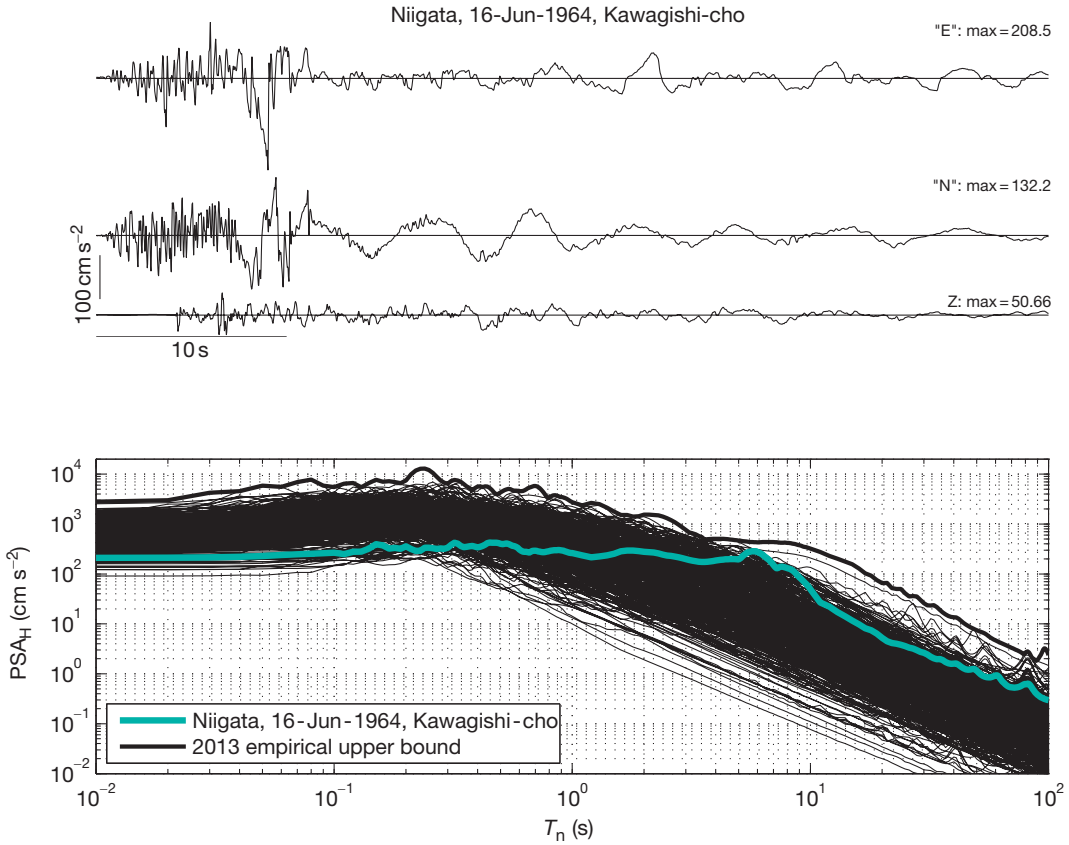


Figure 30 The 1964 Niigata earthquake (M_W 7.6) as recorded at the station in Kawagishi-cho. This digitization is from Kudo et al. (2000). See caption to Figure 4 for more explanation.

times in the record, the curves at 16–18 m depth on the north–south component and at 31 m depth on the east–west component show a stress–strain relationship where μ_{TAN} initially decreases, but then increases with increasing strain. This phenomenon, known as strain hardening, is a characteristic of sandy sites that are saturated with water (e.g., Bonilla et al., 2005; Yang et al., 2002). A second important feature is that for many of the stress–strain curves, μ_{SEC} decreases early in the earthquake, but then does not recover to the earliest values later in the record as the amplitude of shaking decreases. At some depths, it appears to not be fully recovered even in aftershocks up to 6 min later. A third point is that some layers, characterized by different lithologies and documented by Pavlenko and Irikura (2002), behave nearly linearly throughout the earthquake and its aftershocks.

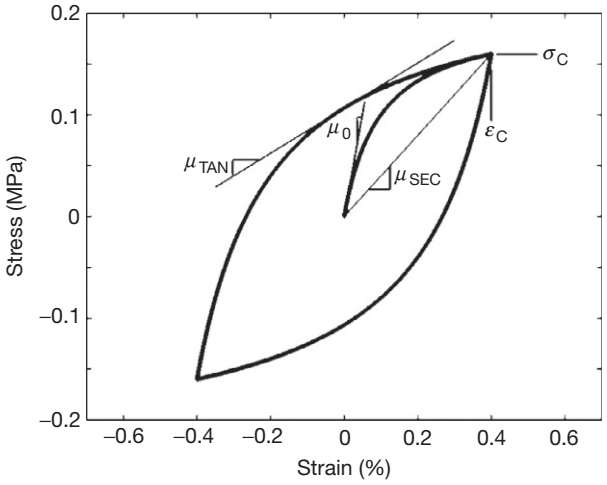


Figure 31 Definition of shear modulus parameters that are used in the studies of nonlinear response of soils. This plot uses the functional form employed by Yu et al. (1993) to show the restoring stresses in a soil as a function of the deformation state during cyclic deformation. The soil starts in an undeformed state, and strain is increased to ε_C, decreased to $-\varepsilon_C$, and then increased again to ε_C, where in this case, $\varepsilon_C = 0.004$.

One empirical model for the average effects of nonlinearity on the amplitudes of ground motions is given in **Tables 6** and **7**. For the strongest level of shaking ($S_s \geq 1.25$ g) in **Table 6**, ground motion amplitudes on site class D are reduced to about 2/3 of what would be expected by simple linear scaling from weak motion data (i.e., $F_a = 1.0$ compared to $F_a = 1.6$ in the weak motion limit). Data from the Port Island site in Kobe suggest that nonlinearity does not reduce the amplitudes of the vertical component of ground motions as much as the amplitude of the horizontal motion, so the ratio of vertical to horizontal peak values might be increased (e.g., Yang and Sato, 2000; Yang et al., 2002).

If μ_{SEC} is decreased during strong motion, as indicated on **Figures 31–33**, the shear velocity should also decrease. Consequently, based on eqn [43], the fundamental resonant frequency of a layer would be decreased. Because of this and also other effects of nonlinearity, the effect of the surface geology, as defined in eqn [58], is expected to be different during strong motion with nonlinear response than it is during weak motion. Furthermore, the increase in ζ or decrease in Q would result in decreased spectral amplitudes at high frequencies. Several studies have recognized those effects (e.g., Aguirre and Irikura, 1997; Beresnev et al., 1998a,b; Chin and Aki, 1991; Cultrera et al., 1999; Frankel et al., 2002a; Higashi and Sasatani, 2000; Rubinstein and Beroza, 2004; Satoh et al., 1995, 2001a; Tsuda et al., 2006b). The largest vector acceleration recorded through the end of 2012 occurred in the 2008 Iwate–Miyagi Nairiku earthquake, recorded at the NIED KiK-

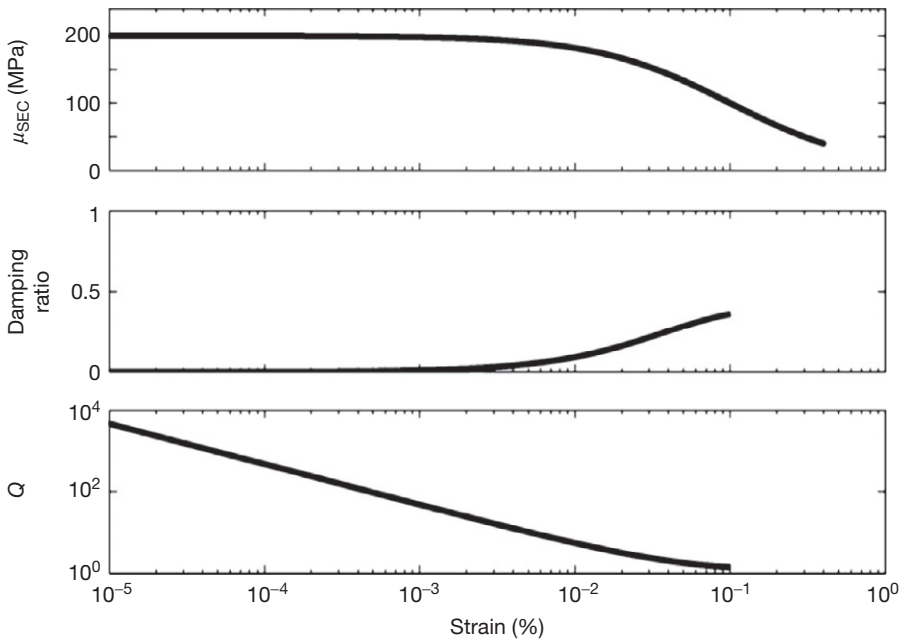


Figure 32 Values of μ_{SEC}, the damping ratio, and Q as a function of ε_C, for the stress–strain curves given in **Figure 31**.

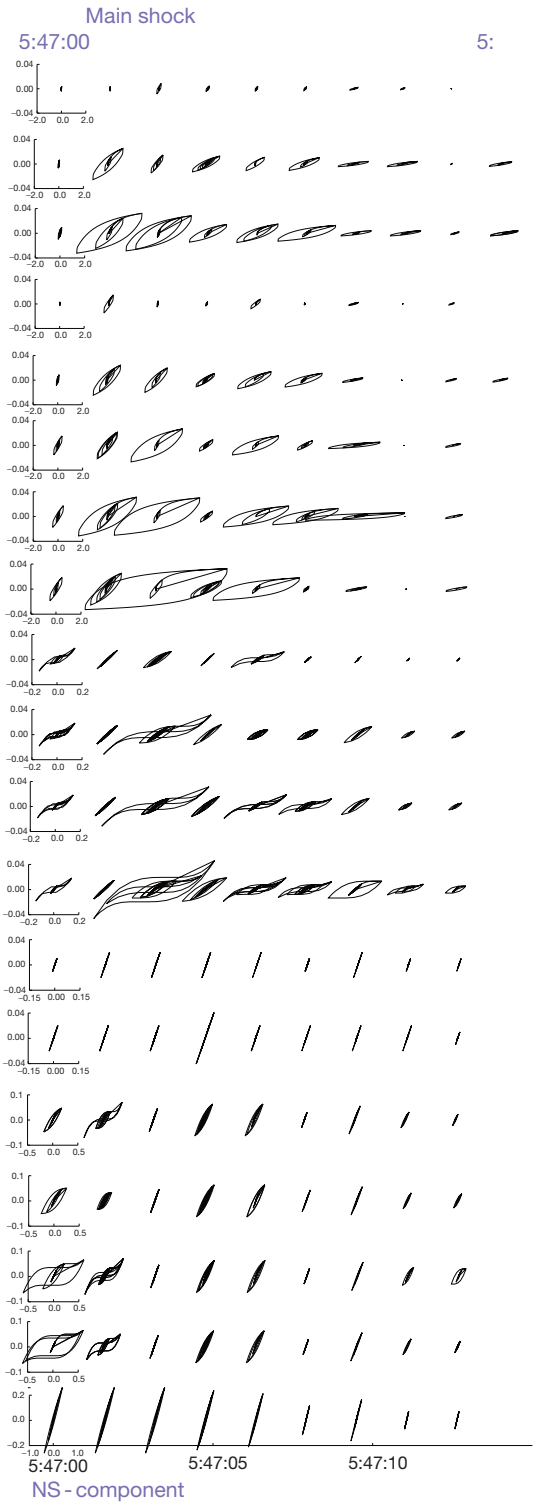


Figure 33 Stress–strain relations, as in **Figure 31**, as a function of depth and time at Port Island during the 1995 Kobe, Japan, earthquake, M_W 6.9, and four of its aftershocks. Each row of this figure represents a different depth, from the surface to 43 m. Main shock columns show stress–strain inverted from ground motion records for different brief time intervals during the main shock. Aftershock columns represent stress–strain inverted from different aftershocks occurring from 2 to 5 min after the main shock. Reproduced from Pavlenko

net station IWTH25. Mogi et al. (2013) demonstrated, using the change in travel time from the downhole to surface sensor at that station, that the shear modulus in the top layer decreased by 60% of its initial value at a strain level of 10^{-3} (0.1%), and a year after the earthquake, it still had not recovered to the initial value. Such a modulus reduction is consistent with **Figure 32**.

The accelerogram at station IWTH25 (**Figure 5**) is notable also for the recognition of the trampoline effect (Aoi et al., 2008). The vertical accelerogram is distinctly asymmetrical, with acceleration pulses in the upward direction being narrow, brief, and substantially stronger than 1 g, while the pulses in the downward direction were broader, flatter, and tending toward peaks of about −1 g. In a borehole instrument 260 m below the surface, the vertical component of acceleration is symmetrical. The interpretation is that during the shaking, the upward accelerations were driven by strong compressional strains in the soil, while the downward accelerations were predominantly controlled by gravity with a secondary contribution from elastic forces in the soil tending to reduce the extensional strain. The observations are consistent with a stress–strain relationship that is nonlinear with the soil much stiffer under compression than under dilation. Yamada et al. (2009) observed that the asymmetrical pulses are similar to a slapdown phase comparable to the observed accelerations of a spalled surface above a nuclear explosion.

Several approaches are available to calculate the effect of nonlinearity on ground motions. In most cases, the overall idea is to start with the ground motions that will be input to the bottom of a flat-layered geologic structure as SH motion and then calculate the response of the structure including nonlinearity. Approaches can be separated into two broad categories: The first is the 'equivalent linear' model and the second is the 'full nonlinear' model. In the equivalent linear approach (e.g., Schnabel et al., 1972), given initial input elastic parameters, one calculates the linear response of the layer (as in **Section 4.20.7.5**). Strains are then calculated, and the secant elastic moduli are determined from the modulus reduction curves (**Figure 32**). These modified elastic moduli are used to recompute the layered response, and the system is iterated until it converges. In the equivalent linear approach, the nonlinearity increases the damping (decreases Q), and thus, high frequencies are suppressed. In the full nonlinear approach (e.g., Archuleta et al., 2003; Elgamal, 1991; Joyner, 1975; Joyner and Chen, 1975; Xu et al., 1999; Yu et al., 1993), the model uses a finite-element solution to the wave equation and follows the stress–strain curves as in **Figure 31**, using the appropriate tangent elastic modulus to determine the stress–strain relationship at each point in the model. In the full nonlinear model, energy is converted to other frequencies (e.g., Yu et al., 1993), and the high-frequency spectrum might increase in some cases. Hartzell et al. (2004) compared approaches to the calculation of nonlinear effects and presented a compact summary of the dependence of modulus reduction curves on material and depth.

O and Irikura K (2002) Changes in shear moduli of liquefied and nonliquefied soils during the 1995 Kobe earthquake and its aftershocks at three vertical-array sites. *Bulletin of the Seismological Society of America* 92(5): 1952–1969.

4.20.7.17 Seismic Site–City Interactions

In placing strong-motion instruments, when the goal is to record 'free-field' motion, seismologists are concerned about the effects of nearby structures. For instance, wind can couple with nearby trees, telephone poles, or other structures to generate undesirable signals on high-gain instruments (e.g., Webb, 2002). For strong motions, man-made structures can similarly modify the ground motion nearby. Jennings (1970) observed that vibrations excited while testing the Millikan Library on the California Institute of Technology campus could be detected up to 10 km away. Wirgin and Bard (1996) and Guéguen et al. (2000, 2002) concluded that the long duration of strong motions in Mexico City from the 1985 Michoacan earthquake (Anderson et al., 1986; **Section 4.20.7.8**) is explained by long-duration vibrations of tall buildings that feed energy back into the ground at an amplitude comparable to the amplitude of ground motions coming from the source. Subsequent studies include Kham et al. (2006), Semblat et al. (2008), and Menglin et al. (2011).

4.20.8 Predicting Strong Ground Motions

4.20.8.1 Synthetic Seismograms

Combining the physics that is summarized in the previous sections, it is possible to generate synthetic ground motions. A vision for strong-motion seismology is that these synthetics should be sufficiently reliable to be used for engineering design of structures. Bolt and Abrahamson (2003) reviewed this topic from the engineering perspective. Another important use is the development of scenarios for possible future earthquakes, to advance preparedness. Faccioli and Pessina (2003) had reviewed this topic from the viewpoint of this application.

The representation theorem, eqn [1], and its expression in the (frequency domain e.g., eqn [4]) provide the mathematical foundation for these approaches. Equation [1] requires two components: the Green's function and a representation of the seismic source. Thus, it would seem that one way to organize this topic would be around these two requirements. In practice, that becomes difficult because many of the methods in current use have intertwined the two in ways that the individual contributions are not intuitively separable. While keeping the representation theorem in mind, this chapter organizes the alternatives in the order of increasing complexity. In general, the more complex methods are more computer-intensive, and the relationships of the components of the more complex models to the representation theorem become more explicit. At this more complex level, where the source and the Green's function are treated as independent modules, there are many possible combinations of source and Green's function.

The models can be evaluated by how well they predict statistical characteristics of the ground motion, or they can be evaluated by how well they are able to match the waveforms of observed strong motions in past earthquakes. To be useful for engineering applications, the statistical characteristics that should be predicted well include Fourier and response spectral amplitudes across the entire frequency band of interest (e.g., 0.1–20 Hz), peak velocity, and duration of shaking in different frequency bands (e.g., Abrahamson et al., 1990; Hartzell et al., 1999, 2005; Schneider et al., 1993). Waveform matching is not expected for engineering applications, since the location of the hypocenter of future large earthquakes is not predictable. However, many but not all of the methods to be discussed aim for waveform matches at low frequencies. Successful waveform matches demonstrate that the particular source model and Green's function combination is capable of predicting actual seismograms – an important feature when the source is a composite kinematic slip function that, while motivated by some physical model for the source, may have a tenuous relationship with physically realizable slip functions (e.g., Zeng and Chen, 2001). Anderson (2004) and Olsen and Mayhew (2010) suggested a score that combines both statistical parameters and waveform matching. As a means of calibrating what a good statistical match might be between synthetic and observed seismograms, Anderson (2004) showed statistics of how well the east and north components of a set of seismograms match each other.

Determining a statistic, or set of statistics, to compare data and predictions is only one part of the problem with evaluating prediction methods. Another challenge comes from calibrating new ideas and determining if they are an improvement over existing methods. To be thorough, a new source model, say, should be compared with an older source model using the same Green's functions, over the set of all of the well-recorded and/or most important strong-motion records. Since a new source model requires the calibration of some model parameters, first, the computational task is enormous and very time consuming. Then, in order to test whether it is reliable for engineering applications, the new method, along with all of the existing methods, should be used to make blind prediction of new data whenever another strong earthquake occurs and is well recorded. In 2012, the Southern California Earthquake Center implemented an efficient system to carry out these functions (Goulet, 2013).

The main approaches can be grouped as methods that use stochastic simulations, that use EGFs, that use synthetic Green's functions with a kinematic source model, and that use synthetic Green's functions with a dynamic source model in a 3D representation of the geologic structure.

4.20.8.1.1 Stochastic simulations

One important approach to generating synthetic seismograms was proposed by Boore (1983). In this approach, one generates a completely random time series, applies a shape function in the time domain so that the time series has a reasonable envelope of amplitude with time, transforms to the frequency domain, modifies the Fourier amplitude spectrum to match the amplitude and shape of a spectral model, and then transforms back to the time domain to obtain the synthetic seismogram. The representation theorem is thus applied in the frequency domain in this approach (eqn [4]). Improvements to this approach are made through improved models of the source spectrum (**Section 4.20.5.3**) and the effects of wave propagation (**Section 4.20.7.15**) to make increasingly sophisticated spectral shapes. Atkinson and Boore (2000) provided a concise summary of why this approach is widely considered useful for engineering seismology applications.

Spectral models may start with point-source approximations, frequently the Brune (1970) model, modified for

attenuation along the path and at the site and for site amplification (e.g., Boore, 1983; Boore and Atkinson, 1987; Boore and Joyner, 1997). Appropriate spectral shapes and the parameters that control them have received considerable attention, especially for regions where observations of strong motions are available not only for seismic hazard analyses (e.g., Atkinson, 1993a,b, 1996, 2004; Atkinson and Boore, 1998; Atkinson and Mereu, 1992; Boatwright and Choy, 1992; Boore, 1986; Boore and Atkinson, 1989, 1992; Edwards and Fäh, 2013; Haddon, 1996, 1997; Rietbrock et al., 2013; Safarshahi et al., 2013) but also for other regions (e.g., Atkinson, 1995; Beresnev and Atkinson, 2002; Margaris and Boore, 1998; Sokolov, 1997). With this calibration, the model has been used for engineering applications, especially to generate synthetic seismograms for a range of magnitudes and distances. The synthetic seismograms have been used as a basis for the development of ground motion prediction equations (GMPEs) (Atkinson, 1984, 1995; Atkinson and Boore, 1995; Atkinson and Somerville, 1994; Boore and Atkinson, 1987).

A finite fault is a more realistic model for larger earthquakes and causes a more complex spectral shape, with multiple corner frequencies (e.g., Atkinson and Cassidy, 2000; Haskell, 1964). Following the conceptual framework given in eqn [3], for larger events close to the fault, Silva et al. (1990) and Atkinson and Silva (1997) summed the contribution of subfaults, where each subfault contribution is treated as a smaller earthquake with energy release timed according to the rupture velocity. Like the point-source model, the stochastic finite-source model has been useful to help develop ground motion predictions (e.g., Atkinson and Silva, 2000) and to interpret well-recorded earthquakes in several regions (e.g., Atkinson and Beresnev, 2002; Atkinson and Boore, 1997; Benetatos and Kiratzi, 2004; Beresnev and Atkinson, 1998, 1999; Gregor et al., 2002; Junn et al., 2002; Roumelioti and Beresnev, 2003; Roumelioti and Kiratzi, 2002; Roumelioti et al., 2004; Silva et al., 2003; Ugurhan et al., 2012).

4.20.8.1.2 Empirical Green's functions

Section 4.20.4.2 showed how eqns [23] and [24] can express the representation theorem, eqn [1], in terms of EGFs. To use an EGF to generate a synthetic for a larger event with moment $M_{0,\text{main}}$, a source time function for the main shock is still needed. Considering that the slip rate $\dot{D}_E$ is limited by the effective shear stress on the fault (eqn [26]), the rise time will be much larger in the main shock than in the EGF. Since the slip function is not known, an alternative is to sum the EGF $N_{EGF}=M_{0,\text{main}}/M_{0,EGF}$ times with appropriate randomized phase delays to simulate the effect of the extended source, guided by the normalizations in eqn [23] or [24]. It is not appropriate to simply multiply the EGF by N_{EGF} because the EGF has a very short rise time. Thus, simple multiplication results in a simulation with the correct moment, but with too high a corner frequency and high-frequency spectral amplitudes that are too large. When more than one EGF is available, the moment release and timing of the different parts of the fault represented by each EGF are incorporated.

Several methods have been proposed to assure that the resulting synthetic seismograms match both the moment and the high-frequency spectral level of the target event (e.g., Dan et al., 1990; Frankel, 1995; Hutchings, 1991, 1994; Irikura, 1983; Joyner and Boore, 1986; Kanamori et al., 1993; Ordaz et al., 1995; Wennerberg, 1990). These methods also deal with the issues that arise when the corner frequency of the EGF falls within the frequency range of interest.

In a strict, formal sense, the need for records from the targeted station limits the applicability of EGFs. Most of the studies cited earlier in the text validate the EGF technique with records from smaller events recorded on the same instrument that recorded a major earthquake. Theoretically, at the future site of a critical engineering structure, a seismic station might be installed with the intention of recording EGFs specifically for that structure. Less formally, a nearby station on similar site conditions could also be used; the uncertainties introduced with this approximation have not been quantified.

The availability of small earthquakes with epicenters on the targeted fault also, in a formal sense, limits the use of the EGF method. Fault segments that rupture in a major earthquake may have few small earthquakes in between the major events. It has long been remarked, for instance, that the segment of the San Andreas Fault in southern California that broke in 1857 has remarkably few small earthquakes located directly on the fault (e.g., Brune and Allen, 1967; Joswig, 2001). To the extent that the near-surface geology has a relatively high influence on the ground motions, as found by Mori (1996), EGFs may not need to come from the targeted fault, although the uncertainty needs to be quantified.

Given the dependence of ground motions on focal mechanism and the direction from the fault to the station, the lack of similarity of the focal mechanism of the main shock and the EGF is potentially limiting. Some studies have suggested that the S-wave radiation pattern breaks down at high frequencies, possibly but not necessarily due to strong scattering. Liu and Helmberger (1985) noted unusual complexity in the radiation pattern of a small earthquake in the Imperial Valley, California. Vidale (1989) found the focal mechanism of the 1987 Whittier Narrows, California, earthquake (M_L 5.9) and its M_L 5.3 aftershock to be important in controlling the distribution of peak acceleration (at frequencies of 3–6 Hz). Satoh (2002) found a transition from amplitudes controlled by the radiation pattern at 3 Hz to isotropic radiation at 6 Hz in the 1998 Miyagiken–Nanbu, Japan, earthquake (M_W 5.0). Pulido and Kubo (2004) found that the transition occurred between 1 and 4 Hz in data from the 2000 Tottori, Japan, earthquake (M_W 6.8). Considering that it is commonly possible to determine the focal mechanism of microearthquakes from high-frequency P-wave arrivals, these differing conclusions are consistent with the idea that scattering more effectively disrupts the mean radiation pattern of larger earthquakes at high frequencies, as later scattered arrivals become more important than direct arrivals. To the extent that is true, the limitations placed on the use of EGF with a different focal mechanism might be muted when simulating ground motions for a large earthquake.

Finally, the EGFs need an adequate signal-to-noise ratio at low frequencies. The importance of this was described in Section 4.20.4.2.1.

4.20.8.1.3 Synthetic Green's functions with kinematic and dynamic sources

Several studies have used synthetic Green's functions to generate synthetic strong ground motion seismograms. The

advantage of these methods is that the seismograms that are generated should be fully consistent with solutions to the wave equation for a distributed source, while the seismograms generated with the previous methods have generally lost information about the phase. Section 4.20.4.1 introduced the main concepts for computing synthetic Green's functions. When these are used, eqn [1] is employed directly as the framework for combining the source and the synthetic Green's functions. Issues in such applications include the choice and validation of the Green's function and the source time function. The use of the finite-difference or finite-element method (reviewed in the succeeding text) to generate Green's functions is preferred at low frequencies. Some studies in this category do not attempt to generate broadband synthetic seismograms, but rather are limited to the frequencies that can be achieved using the finite-difference solutions. Among the studies that do generate broadband seismograms, multiple approaches seem to work to generate the higher frequencies. For instance, Somerville et al. (1991) and Zeng et al. (1994) used Green's functions with a flat-layered velocity model appropriate for the station, which Zeng et al. (1995) modified by introducing a scattering model. Kamae et al. (1998), Hartzell et al. (1999, 2002), Pitarka et al. (2000), and Mena et al. (2006) used a hybrid approach, where the high frequencies are generated by a stochastic model similar to that described in Section 4.20.8.1.1.

A complete review of the finite-element and finite-difference approaches to the generation of synthetic seismograms is beyond the scope of this chapter. The methods themselves are referenced in the articles that report results. The papers by Vidale et al. (1985) and Levander (1988) are among the earliest to use the finite-difference technique to generate synthetic seismograms in a complex geologic structure. Several of the initial applications were to study the effects of a basin on the incoming seismic energy and, with improving computer capabilities, the entire process from the source to the basin (e.g., Frankel, 1993; Frankel and Vidale, 1992; Furumura and Koketsu, 1998; Furumura and Singh, 2002; Graves, 1993, 1998; Inoue and Miyatake, 1997; Nakamura et al., 2012; Olsen and Archuleta, 1996; Olsen et al., 1995a,b, 2006a,b; Pitarka et al., 1998, 2002; Sato et al., 1999; Satoh et al., 2001b; Smerzini and Villani, 2012; Taborda and Bielak, 2013; Wald and Graves, 1998; Yomogida and Etgen, 1993; Yoshimura et al., 2003).

At this level of sophistication, explicit kinematic models for the slip on the source are needed. Several approaches have been applied (e.g., Hartzell et al., 1999, 2005; Irikura, 1983; Tumarkin and Archuleta, 1994; Tumarkin et al., 1994; Zeng et al., 1994). For one approach to describing the source, Somerville et al. (1999) estimated a scaling relation for the size of asperities, in which the asperities are about one-quarter of the total fault rupture area. Miyake et al. (2003) defined the 'strong-motion generation area' as the part of the fault that is the source of large slip velocities, large stress release (~10 MPa), and large amplitudes of high-frequency radiation (0.2–10 Hz). Similarly, Miyake et al. (2003) found that a scaling relation for the rise time obtained by Somerville et al. (1999) is consistent with the rise time of the strong-motion generating area.

To give a sense of how these models might work, we consider briefly a different model, by Zeng et al. (1994). Their kinematic 'composite source model' consists of a large number of subevents, motivated by Frankel (1991). The sizes of the subevents follow a power law distribution, and each subevent is located at random on the fault. Rupture starts at a designated hypocenter and spreads at a designated rupture velocity. Each subevent radiates when the rupture front reaches its center. Figure 34 shows an example of seismograms generated by this method, compared with observations from the 1989 Loma Prieta, California, earthquake. Figure 35 shows an example of the bias and standard deviation achieved by John Anderson for an unpublished model of the 1994 Northridge, California (M_W 6.7), ground motions, in tests described by Goulet (2013). On the left side of Figure 35, one looks for average realizations to show a small bias at all frequencies. On the right side, one looks for the bias to be independent of distance. Other tests considered by Goulet (2013) included generating maps that show the spatial distribution of residuals and comparison of the composite source model and other models with GMPEs (Section 4.20.8.2).

The highest level of complexity in source models is the introduction of dynamic sources (e.g., Aagaard et al., 2001; Beroza and Mikumo, 1996; Duan and Oglesby, 2005; Guatteri et al., 2003, 2004; Miyatake, 2000; Nielsen and Olsen, 2000; Oglesby and Day, 2002; Oglesby et al., 2003, 2004; Olsen et al., 2006a,b; Peyrat and Olsen, 2004). These synthetics draw on current models for the dynamics of slip on the fault. Figure 21 was generated by this approach. These source models do not generate high-frequency energy (e.g., above 1 Hz), but with careful definition of initial stress conditions, they are capable of generating synthetic seismograms that match low frequencies of the observations.

Since the models described in this section follow the physics of the problem closely, one might expect that they would ultimately be the most general methods for ground motion predictions. In the hands of the developers, several of these models have been tuned to give good, unbiased predictions for selected earthquakes. However, the challenge of demonstrating their reliability, as outlined in the introduction of this section, has not yet been fully solved. These models generally have several free source parameters that are available to be adjusted, and blind testing by users who are independent of the model developers has not yet taken place. Their continued development is one of the most exciting frontiers of strong-motion seismology.

4.20.8.2 Ground Motion Prediction Equations

A long-standing goal in engineering seismology has been to develop simple equations that can be used to estimate the amplitudes of various parameters associated with strong ground motions. These models play an essential role in seismic hazard analysis (e.g., Somerville and Moriwaki, 2003). Initially, most such equations attempted to predict peak acceleration as a function of magnitude and distance. With more and higher-quality digital strong-motion data and efficient computer processing, recent models also predict peak velocity and response spectral amplitudes at a range of frequencies and use additional parameters to describe the source, path, and site. As the ability improves to incorporate more physics into the generation of synthetic seismograms, generation of synthetic seismograms may in part supersede the need for GMPEs. However, these equations will remain useful for the foreseeable

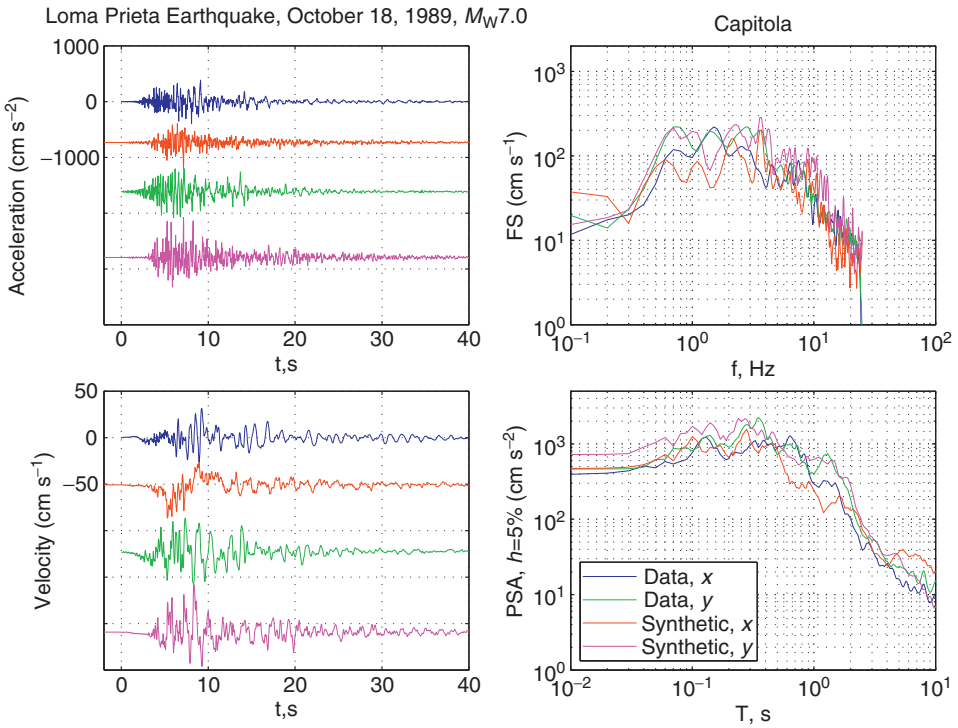


Figure 34 Observed and synthetic seismograms at the Capitola strong motion station recorded in the Loma Prieta earthquake of October 18, 1989 (Mw 7.0). Synthetic seismograms are computed using the composite source model, as implemented by Anderson (2014). Frames show observed and a synthetic realization of acceleration (upper left) and velocity (lower left), and the corresponding Fourier amplitude spectra (upper right) and response spectra (lower right).

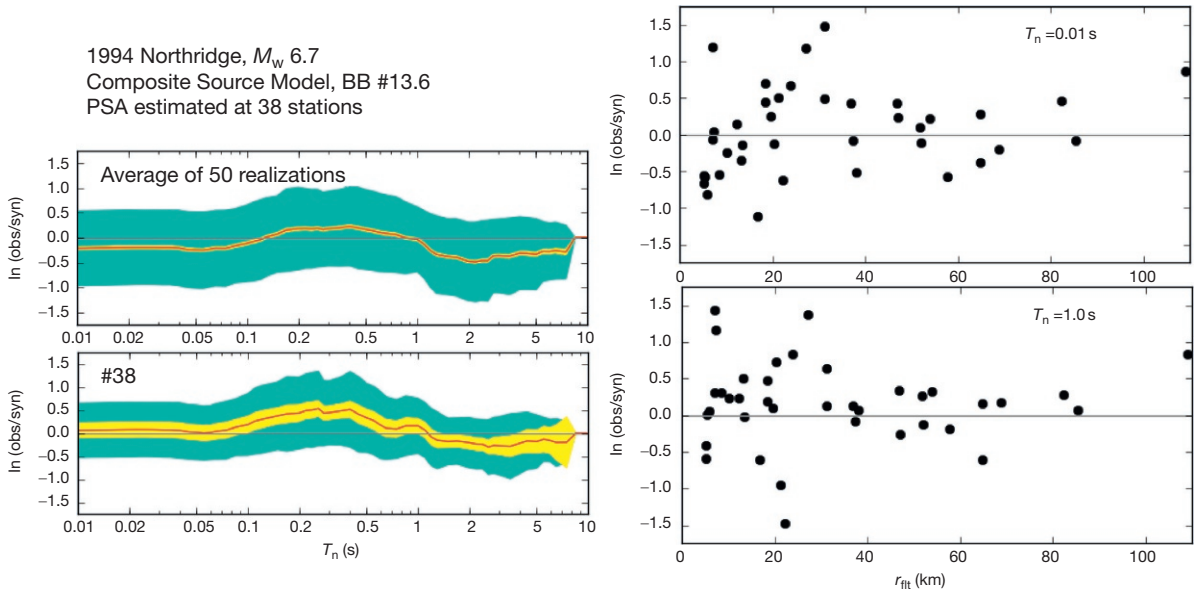


Figure 35 Results of an evaluation of an unpublished implementation by John Anderson of the composite source model (Zeng et al., 1994; **Figure 34**). This implementation was tested in 2013 on the SCEC broadband platform in a test referred to as Broadband Study 13.6. The study used 38 strong-motion recordings from the 1994 Northridge earthquake, M_W 6.7. The test created 50 realizations of the composite source and compared their pseudo-acceleration response spectra (PSA (T_n, ζ)) (ζ = 5% damping) with the data. Bias is defined as the ratio of the observation divided by the prediction of the model. The top left shows the average bias of all 50 realizations and 38 stations. The red central line is the mean bias. The yellow band is the uncertainty in the mean, which is small because there are $50 \times 38 = 1900$ predictions used to find the average. The blue band shows the range of the predictions. The bottom left shows the same information for realization #38, selected as one of the better realizations. The plots on the right show the bias of realization #38 at each of the stations, shown as a function of the nearest distance to the fault. The top right is the bias of PSA (0.01 s, ζ), and the bottom right is the bias of PSA (1.0 s, ζ).

future for probabilistic seismic hazard analyses and for applications that do not require specialized studies.

This chapter refers to these equations as GMPEs. These are sometimes referred to as 'regressions' or 'regression equations,' after the mathematical technique used to estimate the values of the coefficients in the equations. They are also sometimes referred to as 'attenuation equations,' 'attenuation models,' or 'attenuation curves.' Formally, attenuation refers to the rate at which the amplitude of motion decreases with distance. The terminology 'ground motion prediction equations' is more descriptive in defining the purpose of the equations.

Bolt and Abrahamson (2003) and Campbell (2003) reviewed the principles in the development of GMPEs. In the literature, there are large numbers of GMPEs. In particular, Douglas (2011) summarized 289 GMPEs for peak acceleration and 188 for elastic response spectral ordinates. One factor that accounts for the diversity is that whenever significant new data are recorded, it provides the opportunity to improve on past equations. The second factor is that, because of the differences among tectonic environments, many scientists have attempted to develop equations that apply to the region that they are studying, based on only the local data.

Abrahamson and Shedlock (1997) proposed that earthquakes can be divided into three categories for the purposes of development or selection of an appropriate GMPE: shallow earthquakes in active tectonic regions, subduction zones, and stable continental interiors. Support for their viewpoint would be based on an appeal to fundamental physics. Considering that the chemical composition of the continental crust and the thermal structure of the crust are not that variable, it is difficult to see why the physics of slip on a fault or wave propagation should be different depending on the continent. The physical argument for separating stable continental interiors from active tectonic regions would be based on the presence of much older, cooler, thicker, and less fractured crust in the stable regions, resulting in different wave-propagation and attenuation characteristics.

There are some common features to GMPEs. Following the notation of Anderson and Uchiyama (2011), let Y_{es} represent any of the parameters used to measure the amplitude of ground motion, as observed at a station at location x_s from an earthquake at location x_e. GMPEs generally take the form

$$\ln Y_{es} = f(Z_{es}, \Theta) + \delta^{I}_{es} \qquad [62]$$

In eqn [62], $f(Z_{es}, \Theta)$ is an equation for the log of the mean amplitude of ground motion, where Z_{es} is a vector of the predictor variables, Θ is a vector of parameters determined by regression from the data, and δ^{I}_{es} is the residual for this datum. In developing the equation, δ^{I}_{es} is generally treated as a random variable that belongs to a normal distribution with zero mean. The uncertainty in the model is the standard deviation of δ^{I}_{es}:

$$\sigma^{I} = \sqrt{\frac{1}{n-1}\sum_{e,s}^{n}\left(\ln Y_{ex} - \ln \hat{Y}^{I}_{es}\right)^{2}} \qquad [63]$$

In eqn [63],

$$\ln \hat{Y}^{I}_{es} = f(Z_{es}, \Theta) \qquad [64]$$

is the GMPE. The sum in eqn [63] is over all n available combinations of e and s.

Typical parameters represented by Y_{es} are peak acceleration, peak velocity, peak displacement, and pseudo-acceleration response spectra with 5% damping for periods from 10 to 0.01 s (100 Hz). Because of the asymptotic properties of the response spectrum, at a frequency of 100 Hz, the response spectrum is generally identical to peak acceleration; determining the response spectrum at oscillator frequencies/periods outside of the range of frequencies with significant amplitudes on the Fourier amplitude spectrum of the accelerogram is meaningful.

Predictor variables in eqn [64] invariably include M_W. Functional form for $f(Z_{es}, \Theta)$ in eqn [62] generally includes a linear term in M_W and often some additional magnitude dependence terms that allow Y_{es} to 'saturate,' that is, reach a maximum value, as M_W increases. Some older models use a different magnitude scale, but M_W is now almost universally preferred. The distance to the nearest point on the extended fault, r_{flt}, is usually another predictor variable. While r_{flt} is not universally accepted as the best way to measure the distance to an extended fault, there is agreement that the distance to the hypocenter or epicenter is not appropriate. Noting that eqns [9] and [10] have distance terms in the form of r^{-1}, that taking the log of eqns [9] and [10] will result in terms of the form $-\ln r$, that wave propagation is more complex in a layered medium than in a homogeneous space, and that the finite-fault size affects the rate of ground motion attenuation, many models include terms in the form $-b_r \ln r_{flt}$. This form sets the distance dependence of geometric spreading to an unknown parameter that is estimated by regression.

Several additional variables are generally included. The site condition was often entered as a categorical variable in earlier models (e.g., Abrahamson and Silva, 1997). With improved data about the site conditions, several models (e.g., Abrahamson and Silva, 2008; Boore and Atkinson, 2008; Campbell and Bozorgnia, 2008; Chiou and Youngs, 2008) now use V_{S30}, thus aligning the equations with the NEHRP classification (**Table 5**). Alternative site classification includes the site period (e.g., Di Alessandro et al., 2012; Zhao and Xu, 2013). Other variables include an adjustment for nonlinearity, an adjustment for geometries where the station is on the hanging wall of a fault, an adjustment for directivity, and fault type (strike-slip, normal, or thrust), generally entered as a categorical variable.

The error term requires particular attention. Averaged over all stations for a single event, one obtains event terms, E_e. Averaged over all events for a single station, one obtains station terms, S_s. Event terms are expected be normally distributed with zero mean and with standard deviations of say σ_{Events}. Station terms may not be normally distributed, since a mechanism for them to have infinite tails is difficult to identify (e.g., Anderson and Brune, 1999a), but the central part of the distribution may be approximately normal with standard deviation $\sigma_{Stations}$. For a random future earthquake, when neither the station nor the source term is known, the total standard deviation is $\sigma^{I} = [\sigma^2_{Events} + \sigma^2_{Stations} + (\sigma^{III})^2]^{1/2}$, where σ^{III} is found from the equivalent of eqn [62], after adjusting for the event and station terms. Overestimates of σ^{I} can have serious negative consequences for probabilistic seismic hazard estimates at low probabilities, so several recent studies have focused on correct estimation of the uncertainties. The station term, at

least, can be estimated in advance allowing estimates of single-station sigma, $\sigma^{II}=[\sigma^2_{Events}+(\sigma^{III})^2]^{1/2}$. These are significantly smaller than σ^{I} (e.g., Anderson and Uchiyama, 2011; Rodriguez-Marek et al., 2011).

One notable characteristic about the process of developing the GMPE is the application of a team approach. For example, in the Next Generation Attenuation project, the development of the database of ground motion observations for the modeling teams is carried out as a separate project (Chiou et al., 2008). This database has drawn on a data set of source models for major earthquakes (Mai et al., 2006). Synthetic seismograms (**Section 4.20.8.1**) were developed to provide theoretical guidance for the modeling as desired. Five teams of modelers were formed to draw upon this data. Each team generates its own GMPE (Abrahamson and Silva, 2008; Boore and Atkinson, 2008; Campbell and Bozorgnia, 2008; Chiou and Youngs, 2008; Idriss, 2008), but the teams interacted considerably. Finally, throughout the project, there have been frequent open meetings that provide for interaction among the NGA project team, scientists, and users (Power et al., 2008).

Figure 36 shows the predictions of the Abrahamson and Silva (2008) model for peak acceleration and associated standard deviation for three magnitudes on a site characterized by $V=760\ m\,s^{-1}$. **Figure 37** shows the corresponding model for peak velocity, and **Figure 38** shows predicted acceleration response spectra. **Figures 36–38** show that total standard deviations (σ^{I}) in the prediction of ground motion are in the range from 0.5 to 0.8 natural logarithm units. Since this is a lognormal distribution, a standard deviation of 0.69 corresponds to a multiplicative factor of 2.

Figures 36–38 show ground motions increasing less as magnitude increases from 6.5 to 7.5 than it does as magnitude increases from 5.5 to 6.5. This tendency is referred to as saturation. The tendency toward saturation is more pronounced for higher-frequency or shorter-period ground motion parameters. Thus, saturation is more prominent for peak acceleration than for peak velocity and more prominent at short periods than at long periods on **Figure 38**. As noted in **Section 4.20.5.4**, large magnitudes are correlated with large faults. However, ground motion near a site depends more on the nature of fault slip near that site than on how far beyond that site the fault ruptures. Thus, if there are limits to the stresses that are accumulated and released on a local segment of a fault, those are expected to dominate at short distances, leading to the observed tendency toward saturation with magnitude.

Figures 36 and **37** also show the attenuation of peak amplitudes to be faster for small magnitudes than for large magnitudes. Anderson and Quaas (1988) showed a good example of how the amplitude of peak acceleration decreases much more slowly as distance increases for large magnitude earthquakes than for small magnitude earthquakes. Rogers and Perkins (1996) and Anderson (2000a) presented differing explanations for this phenomenon. Rogers and Perkins pointed out that when more parts of the fault contribute, the peak acceleration will follow the statistics of extreme values. Anderson pointed out that for the larger fault, with a long source duration, the late arrivals from parts of the fault that rupture first arrive at the same time as the earlier arrivals from later-rupturing parts of the fault. Both phenomena would contribute. Neither study, however, mentions nonlinear soil response, which can depress the amplitude of high-frequency motions near the source but not at large distances.

It will be noted that the median ground motions in **Figures 36–38** are much smaller than the highest recorded motions shown in **Figures 2** and **3**. The magnitudes of most of these records are smaller than M_W 7.5 (**Figure 2**), so the explanation is not that the magnitude used to predict motions in **Figures 36–38** is too small. Part of the explanation is that these models are for a site with $V_{S30}=760\ m\,s^{-1}$, which is a rather stiff site (**Section 4.20.7.1**). Since many strong-motion records are recorded on sediments in site classes C, D, and E, somewhat higher amplitudes are expected (e.g., **Section 4.20.7.2**). Many of the high-amplitude records also show strong resonant peaks and are affected by other near-surface

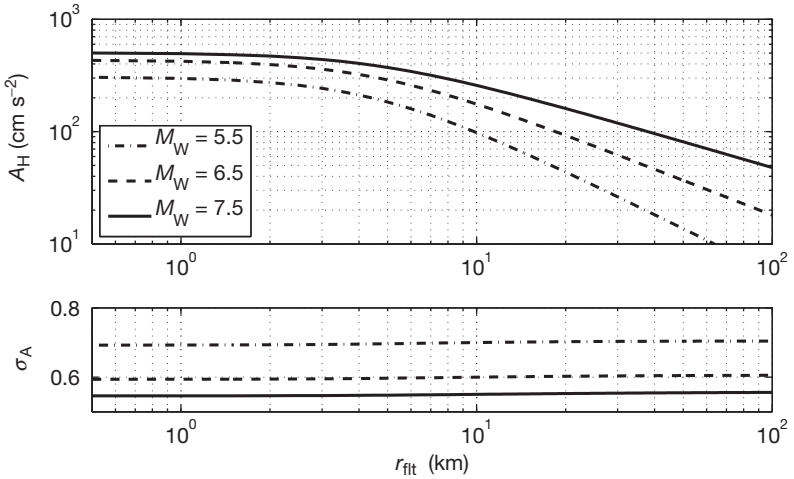


Figure 36 (Top) Median estimates of peak ground acceleration as a function of distance to the nearest point on the fault, in the model of Abrahamson and Silva (2008). These ground motion predictions are for strike-slip earthquakes in the continental crust, for moment magnitudes of 5.5–7.5, and for near-surface geology characterized by $V_{S30}=760\ m\,s^{-1}$. (Bottom) Standard deviation, σ_A, of the distribution of the data minus predictions for natural log of peak acceleration, as a function of M_W and distance, in the model of Abrahamson and Silva (2008).

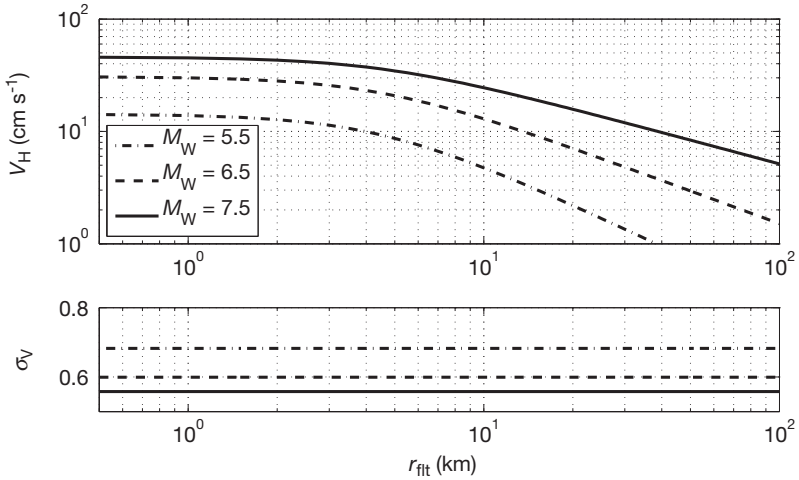


Figure 37 (Top) Median estimates of peak ground velocity as a function of distance to the nearest point on the fault, in the model of Abrahamson and Silva (2008). These ground motion predictions are for strike-slip earthquakes in the continental crust, for moment magnitudes of 5.5–7.5, and for near-surface geology characterized by $V_{S30}=760$ m s^{-1}. (Bottom) Standard deviation, σ_V, of the distribution of the data minus predictions for natural log of peak velocity, as a function of M_W and distance, in the model of Abrahamson and Silva (2008).

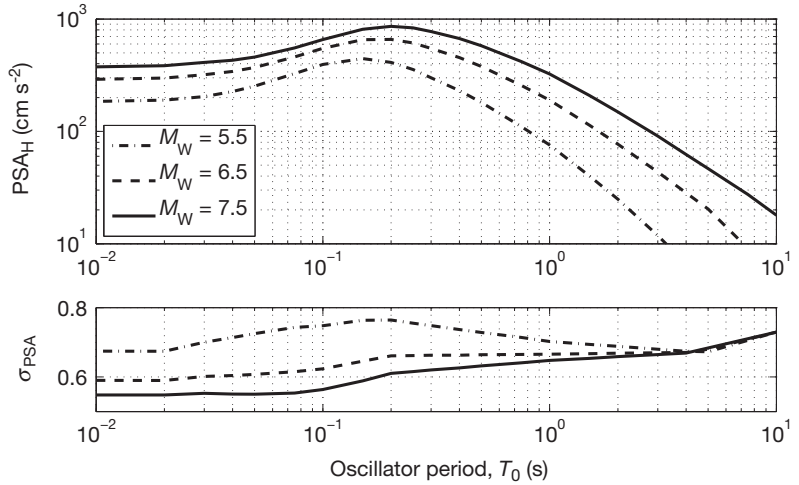


Figure 38 Ground motion predictions of Abrahamson and Silva (2008) for pseudo-acceleration response spectra (PSA) with 5% damping. These predictions are for a location at distance $r_{flt}=5$ km from a vertical strike-slip fault and for near-surface geology characterized by $V_{S30}=760$ m s^{-1}. The acceleration response at the oscillator period T_0 gives the peak response of a damped single-degree-of-freedom oscillator to the accelerogram. Because of its usefulness for approximating the response of a simple structure, PSA is commonly used in engineering seismology instead of the Fourier spectrum (FS). At short periods or high frequencies ($f_0=1/T_0$), PSA is asymptotic to peak acceleration, in contrast to the Fourier amplitude spectrum that falls off rapidly. At low frequencies, the asymptote is controlled by the peak displacement on the integrated accelerogram. At frequencies between those controlled by the asymptotes, PSA is related to $(2\pi f_0)$ FS (f_0) (Udwadia and Trifunac, 1973). At the bottom, the figure shows the standard deviation of the prediction equations, σ_{PSA}, in natural logarithm units. In this model, the standard deviation is a function of both period and magnitude.

effects described in **Section 4.20.7**. Because every site is unique and the list of predictor variables in eqn [62] is short, there will inevitably be residual misfits at all of the observations modeled by eqn [62]. The residuals appear to be normally distributed out to at least 2σ, and residuals that are over 3σ from the mean are observed (Bommer et al., 2004).

If the residuals δ_{es} are normally distributed, then with longer observation times, arbitrarily large amplitudes of ground motion should be observed. The issue of 'extreme ground motions' is whether it is possible to place a practical upper bound on earthquake motions. For instance, the 1996 and 2002 National Seismic Hazard Maps for the United States (Frankel et al., 2002b; Petersen et al., 2008) show levels of ground motions close to the San Andreas Fault that Brune (1996) had argued are not consistent with observations of precariously balanced rocks. Several subsequent studies have strengthened the scientific basis for the use of precarious rocks to test hazard curves at low probabilities (e.g., Anderson

et al., 2000a, 2011; Anooshehpoor and Brune, 2002; Anooshehpoor et al., 1999, 2004, 2013; Bell et al., 1998; Briggs et al., 2013; Brune, 2000, 2001, 2002, 2003; Brune et al., 2003, 2004, 2005, 2006; Purvance, 2005; Shi et al., 1996; Stirling et al., 2002). Anderson and Brune (1999b) also found inconsistencies with the national map in Nevada. One possible explanation is that the exceptions might be the result of uncertainties in the GMPE. Anderson and Brune (1999a) suggested that the cause is an ergodic assumption in the probabilistic seismic hazard analysis methodology: The uncertainties in eqn [50] include spatial uncertainties in the ground motions, but in a probabilistic seismic hazard analysis (PSHA), the type of uncertainty that is needed is the uncertainty of the ground motion for an individual site.

Research is ongoing to improve GMPEs and also to find ways to use GMPEs developed in areas with relatively abundant data in regions where data are sparse (e.g., Arango et al., 2012; Beauval et al., 2012; Delavaud et al., 2012).

4.20.9 Conclusions

The global set of strong-motion data is still very sparse for some critical situations. Thus, continuing to gather data is essential. One situation is ground motions close to the fault in large magnitude earthquakes in the continental crust. Another is observations in continental interiors where Q is large and ground motions are potentially damaging to large distances. Continental earthquakes with $M > 7$ are relatively rare events, as pointed out by Anderson et al. (2005b), and every time one occurs outside of a strong-motion network, an improved capability to make informed decisions on how to build to protect people from earthquakes has most likely been deferred for a few more years. Modern strong-motion networks serve additional critical purposes: They rapidly inform emergency response personnel about where shaking in an earthquake was severe. Strong-motion networks also provide estimates of motion at sites where structures fail, thus allowing better understanding of the causes of failures and improved future construction. Thus, expanded strong-motion instrumentation is strongly recommended.

It is appropriate to return to the questions raised at the beginning of this chapter: Is the physics of strong ground motion understood, and is it possible to estimate the amplitudes of ground motions well enough for engineering applications? Addressing the second question first, I conclude that although uncertainties are large where data are sparse, we have a reasonably good empirical understanding of the amplitudes of strong ground motions. Thus, while there is much to be done to reduce uncertainties, it is safe to say that most earthquake disasters do not occur because of a lack of knowledge of the ground motions. A very small fraction of the world's large earthquakes occur in places where they are not expected, potentially causing a disaster. But most earthquake disasters occur because existing techniques to identify sources and estimate amplitudes of probable ground motions and existing knowledge of how to construct structures to resist those motions are not being applied due to poverty, priorities, or politics.

Strong motions with realistic appearances and approximately correct amplitudes can be synthesized using sound mathematical principles, a variety of calibrated source models, and models for wave propagation in simplified models for Earth structure. The kinematic and dynamic source models cannot be tested more directly because the source is not observable. With that caveat, most of the phenomena observed to date on strong-motion records can be explained on the basis of well-understood wave-propagation processes acting on the waves radiated from the source.

Appendix Sources of Strong-Motion Data, 2012

Organization	*Websites*
Global: National Geophysical Data Center (NGDC) of the National Oceanic and Atmospheric Administration (NOAA) (the United States). Significant records worldwide 1933–94	http://www.ngdc.noaa.gov/hazard/strong.shtml
Global: Consortium of Organizations of Strong Motion Observation Systems (COSMOS) (the United States, selected global, through ~2008)	http://www.cosmos-eq.org/
Global: Pacific Earthquake Engineering Research Center (PEER) Strong Motion Database (selected significant worldwide earthquakes, uniformly processed)	http://peer.berkeley.edu/smcat/
Europe: European Strong-Motion Database (ESD)	http://www.isesd.hi.is/ESD_Local/frameset.htm
Europe: Quakeportal (links to several online strong-motion databases)	http://www.quakeportal.eu/strong-motion-databases.html
Canada: Canadian National Seismograph Network (CNSN)	http://www.earthquakescanada.nrcan.gc.ca/stnsdata/cnsn/
Chile: Laboratorio de Sismología de Movimiento Fuerte, Escuela de Ingeniería, Pontificia Universidad Católica de Chile	http://www2.ing.puc.cl/~wwwice/sismologia/
China: China Strong Motion Network Center (CSMNC)	http://www.csmnc.net/
Greece: Institute of Engineering Seismology and Earthquake Engineering Research and Technical Institute (ITSAK)	http://www.itsak.gr/en/page/networks/acc_network/
Iran: Iran Strong Motion Network (ISMN)	http://www.bhrc.ac.ir/portal/Default.aspx?tabid=635
Italy: Italian Accelerometric Archive (ITACA)	http://itaca.mi.ingv.it/ItacaNet/
Japan: National Research Institute for Earth Science and Disaster Prevention (NIED)	http://www.kik.bosai.go.jp/ http://www.kyoshin.bosai.go.jp/
Japan: Building Research Institute (BRI)	http://smo.kenken.go.jp/
New Zealand: GeoNet Strong-Motion Data	http://info.geonet.org.nz/display/appdata/Strong-Motion+Data
Taiwan: Institute of Earth Sciences (IES), Academia Sinica, Data Management Center for Strong Motion Seismology	http://www.earth.sinica.edu.tw/~smdmc/

Turkey: TR-NSMN: Disaster and Emergency Management Presidency Earthquake Department	http://kyh.deprem.gov.tr/indexen.htm
United States: CESMD: Center for Engineering Strong Motion Data (CESMD) (current United States, selected global)	http://www.strongmotioncenter.org/
United States: National Strong-Motion Project (NSMP), US Geological Survey (current United States, selected North America)	http://nsmp.wr.usgs.gov/

Acknowledgments

The field work needed to gather strong-motion data requires great patience and perseverance, since significant strong-motion records are even less common than significant earthquakes. The efforts of hundreds of people and support of their organizations to obtain this valuable data are gratefully acknowledged.

References

Aagaard BT, Heaton TH, and Hall JF (2001) Dynamic earthquake ruptures in the presence of lithostatic normal stresses: Implications for friction models and heat production. *Bulletin of the Seismological Society of America* 91: 1765–1796.

Abercrombie R (1995) Earthquake source scaling relationships from −1 to 5 ML using seismograms recorded at 2.5 km depth. *Journal of Geophysical Research* 100: 24,015–24,036.

Abrahamson NA, Schneider JF, and Stepp JC (1991) Empirical spatial coherency functions for application to soil-structure interaction analysis. *Earthquake Spectra* 7: 1–27.

Abrahamson NA and Shedlock KM (1997) Overview. *Seismological Research Letters* 67: 9–23.

Abrahamson NA and Silva W (1997) Empirical response spectral attenuation relations for shallow crustal earthquakes. *Seismological Research Letters* 68: 94–127.

Abrahamson NA and Silva W (2008) Summary of the Abrahamson & Silva NGA ground-motion relations. *Earthquake Spectra* 24: 67–97.

Abrahamson NA, Somerville PG, and Cornell CA (1990) Uncertainty in numerical strong motion predictions. In: *Proceedings of the Fourth U.S. National Conference on Earthquake Engineering*, vol. 1, pp. 407–416.

Aguirre J and Irikura K (1997) Nonlinearity, liquefaction, and velocity variation of soft soil layers in Port Island, Kobe, during the Hyogoken Nanbu earthquake. *Bulletin of the Seismological Society of America* 87: 1244–1258.

Aki K (1967) Scaling law of seismic spectrum. *Journal of Geophysical Research* 72(4): 1217–1231.

Aki K (1969) Analysis of the seismic coda of local earthquakes as scattered waves. *Journal of Geophysical Research* 74: 615–631.

Aki K (1980) Attenuation of shear-waves in the lithosphere for frequencies from 0.05 to 25 Hz. *Physics of the Earth and Planetary Interiors* 21: 50–60.

Aki K (1992) Scattering conversion P to S vs. S to P. *Bulletin of the Seismological Society of America* 82: 1969–1972.

Aki K and Chouet B (1975) Origin of coda waves: Source, attenuation and scattering effects. *Journal of Geophysical Research* 89: 3322–3342.

Aki K and Richards PG (1980) *Quantitative Seismology Theory and Methods*. San Francisco, CA: W. H. Freeman and Co.

Aki K and Richards PG (2002) *Quantitative Seismology*, 2nd edn. Sausalito, CA: University Science Books, 700 pp.

Akinci A, Malagnini L, Herrmann RB, Pino NA, Scognamiglio L, and Eyidogan H (2001) High-frequency ground motion in the Erzincan region, Turkey: Inferences from small earthquakes. *Bulletin of the Seismological Society of America* 91: 1446–1455.

Allmann BP and Shearer PM (2009) Global variations of stress drop for moderate to large earthquakes. *Journal of Geophysical Research* 114: B01310. http://dx.doi.org/10.1029/2008JB005821.

Anderson JG (1986) Implication of attenuation for studies of the earthquake source. *Earthquake Source Mechanics, Geophysical Monograph 37. Maurice Ewing Series 6*, pp. 311–318. Washington, DC: American Geophysical Union.

Anderson JG (2000) Expected shape of regressions for ground-motion parameters on rock. *Bulletin of the Seismological Society of America* 90: S43–S52.

Anderson JG (2003) Strong-motion seismology. In: Lee WHK, Kanamori H, and Jennings PC (eds.) *International Handbook of Earthquake and Engineering Seismology*, vol. 81B, pp. 937–965. Amsterdam: International Association of Seismology and Physics of the Earth's Interior, Committee on Education, Academic Press.

Anderson JG (2004) Quantitative measure of the goodness-of-fit of synthetic seismograms. In: *Proceedings, 13th World Conference on Earthquake Engineering*, Vancouver, BC, 1–6 August 2004, Paper No. 243.

Anderson JG (2010) Source and site characteristics of earthquakes that have caused exceptional ground accelerations and velocities. *Bulletin of the Seismological Society of America* 100(1): 1–36. http://dx.doi.org/10.1785/0120080375.

Anderson JG (2012) An overview of the largest amplitudes in recorded ground motions. In: *Proceedings of the 15th World Conference on Earthquake Engineering*, Lisbon, Portugal, 24–28 September 2012, Paper 3911.

Anderson JG (2013) Surface motions on near-distance rock sites in the 2011 Tohoku earthquake, G. *Earthquake Spectra* 29(S1): S23–S35.

Anderson JG, Bodin P, Brune J, Prince J, and Singh SK (1986) Strong ground motion and source mechanism of the Mexico earthquake of September 19, 1985. *Science* 233: 1043–1049.

Anderson JG and Brune JN (1999a) Probabilistic seismic hazard analysis without the ergodic assumption. *Seismological Research Letters* 70: 19–28.

Anderson JG and Brune JN (1999b) Methodology for using precarious rocks in Nevada to test seismic hazard models. *Bulletin of the Seismological Society of America* 89: 456–467.

Anderson JG, Brune JN, Anooshehpoor R, and Ni S-D (2000a) New ground motion data and concepts in seismic hazard analysis. *Current Science* 79(9): 1278–1290.

Anderson JG, Brune JN, Biasi G, Anooshehpoor A, and Purvance M (2011) Workshop report: Applications of precarious rocks and related fragile geological features to U.S. national hazard maps. *Seismological Research Letters* 82: 431–441.

Anderson JG, Brune JN, Purvance M, and Anooshehpoor A (2005b) Data needs for improved probabilistic seismic hazard analysis. In: Gulkan P and Anderson JG (eds.) *Directions in Strong Motion Instrumentation. NATO Science Series, IV: Earth and Environmental Sciences*, vol. 58, pp. 1–24. Dordrecht, The Netherlands: Springer.

Anderson JG and Hough S (1984) A model for the shape of the Fourier amplitude spectrum of acceleration at high frequencies. *Bulletin of the Seismological Society of America* 74: 1969–1994.

Anderson JG and Humphrey JR (1991) A least squares method for objective determination of earthquake source parameters. *Seismological Research Letters* 62: 201–209.

Anderson JG, Koketsu K, and Miyake H (2013) Expanded compilation and statistics of exceptional ground motions. *Final Technical Report for US Geological Survey*, Award No. G12AP20024, Award Year 2012, http://earthquake.usgs.gov/research/external/research.php, 57 pages.

Anderson JG and Luco JE (1983a) Parametric study of near-field ground motion for a strike slip dislocation model. *Bulletin of the Seismological Society of America* 73: 23–43.

Anderson JG and Luco JE (1983b) Parametric study of near field ground motions for oblique-slip and dip-slip dislocation models. *Bulletin of the Seismological Society of America* 73: 45–57.

Anderson JG and Quaas R (1988) Effect of magnitude on the character of strong ground motion: An example from the Guerrero, Mexico, strong motion network. *Earthquake Spectra* 4: 635–646.

Anderson JG, Sucuoglu H, Erbernik A, et al. (2000b). Implications for seismic hazard analysis. In: Youd TL, Bardet JP, and Bray JD (eds.), Kocaeli, Turkey, Earthquake of August 17, 1999, Reconnaissance Report, *Earthquake Spectra* 16 (supplement A): 113–137.

Anderson JG and Uchiyama Y (2011) A methodology to improve ground-motion prediction equations by including path corrections. *Bulletin of the Seismological Society of America* 101(4): 1822–1846.

Anderson JG, Wesnousky SG, and Stirling MW (1996) Earthquake size as a function of fault slip rate. *Bulletin of the Seismological Society of America* 86: 683–690.

Anderson JG and Yu G (1996) Predictability of strong motions from the Northridge, California, earthquake. *Bulletin of the Seismological Society of America* 86: S100–S114.

Anderson JG, Purvance M, Brune JN, and Anooshehpoor A (2005a) Developing Constraints on extreme ground motions based on seismic and geological observations (extended abstract). In: *2nd International Workshop, Strong Ground*

Motion Prediction and Earthquake Tectonics in Urban Areas, 25–27 October 2005, pp. 63–66. Japan: Earthquake Research Institute, University of Tokyo.

Anooshehpoor A and Brune JN (1989) Foam rubber modeling of topographic and dam interaction effects at Pacoima Dam. *Bulletin of the Seismological Society of America* 79(5): 1347–1360.

Anooshehpoor A and Brune JN (2002) Verification of precarious rock methodology using shake table tests of rock models. *Soil Dynamics and Earthquake Engineering* 22: 917–922.

Anooshehpoor A, Brune JN, Daemen J, and Purvance MD (2013) Constraints on ground accelerations inferred from unfractured hoodoos near the Garlock Fault, California. *Bulletin of the Seismological Society of America* 103(1): 99–106.

Anooshehpoor A, Brune JN, and Zeng Y (2004) Methodology for obtaining constraints on ground motion from precariously balanced rocks. *Bulletin of the Seismological Society of America* 94: 285–303.

Anooshehpoor A, Heaton TH, Shi BP, and Brune JN (1999) Estimates of the ground accelerations at Point Reyes Station during the 1906 San Francisco earthquake. *Bulletin of the Seismological Society of America* 89(4): 845–853.

Aoi S, Kunugi T, and Fujiwara H (2008) Trampoline effect in extreme ground motion. *Science* 322: 727–730.

Apsel RJ and Luco JE (1983) On the Green's functions for a layered half space: Part II. *Bulletin of the Seismological Society of America* 73: 931–951.

Arango MC, Strasser FO, Bommer JJ, et al. (2012) An evaluation of the applicability of current ground-motion models to the South and Central American subduction zones. *Bulletin of the Seismological Society of America* 102(1): 143–168.

Archuleta RJ, Cranswick E, Mueller C, and Spudich P (1982) Source parameters of the 1980 Mammoth Lakes. California, earthquake sequence. *Journal of Geophysical Research* 87: 4595–4607.

Archuleta RJ and Hartzell SH (1981) Effects of fault finiteness on near-source ground motion. *Bulletin of the Seismological Society of America* 71: 939–957.

Archuleta RJ, Liu P, Steidl JH, Bonilla LF, Lavallée D, and Heuze F (2003) Finite-fault site-specific acceleration time histories that include nonlinear soil response. *Physics of the Earth and Planetary Interiors* 137: 153–181.

Asano K and Iwata T (2011) Strong ground motion generation during the 2011 Tohoku-Oki Earthquake. Abstract, 2011 Fall Annual Meeting, American Geophysical Union, #U42A-03.

Assimaki D and Gazetas G (2004) Soil and topographic amplification on canyon banks and the 1999 Athens earthquake. *Journal of Earthquake Engineering* 8: 1–43.

Assimaki D, Gazetas G, and Kausel E (2005) Effects of local soil conditions on the topographic aggravation of seismic motions: Parametric investigation and recorded field evidence from the 1999 Athens earthquake. *Bulletin of the Seismological Society of America* 95: 1059–1089.

Assimaki D, Li W, Steidl JH, and Tsuda K (2008) Site amplification and attenuation via downhole array seismogram inversion: a comparative study of the 2003 Miyagi-Oki aftershock sequence. *Bulletin of the Seismological Society of America* 98: 301–330. http://dx.doi.org/10.1785/0120070030.

Atkinson GM (1984) Attenuation of strong ground motion in Canada from a random vibrations approach. *Bulletin of the Seismological Society of America* 74(6): 2629–2653.

Atkinson G (1993a) H/V ratio for hard-rock sites. *Bulletin of the Seismological Society of America* 83: 587–596.

Atkinson G (1993b) Source spectra for earthquakes in eastern North America. *Bulletin of the Seismological Society of America* 83: 1778–1798.

Atkinson GM (1995) Attenuation and source parameters of earthquakes in the Cascadia region. *Bulletin of the Seismological Society of America* 85(5): 1327–1342.

Atkinson GM (1996) The high-frequency shape of the source spectrum for earthquakes in Eastern and Western Canada. *Bulletin of the Seismological Society of America* 86(1A): 106–112.

Atkinson GM (2004) Empirical attenuation of ground-motion spectral amplitudes in southeastern Canada and the northeastern United States. *Bulletin of the Seismological Society of America* 94(3): 1079–1095.

Atkinson GM and Beresnev IA (2002) Ground motions at Memphis and St. Louis from M 7.5–8.0 earthquakes in the New Madrid seismic zone. *Bulletin of the Seismological Society of America* 92(3): 1015–1024.

Atkinson GM and Boore DM (1995) Ground-motion relations for eastern North America. *Bulletin of the Seismological Society of America* 85: 17–30.

Atkinson GM and Boore DM (1997) Stochastic point-source modeling of ground motions in the Cascadia region. *Seismological Research Letters* 67: 74–85.

Atkinson GM and Boore DM (1998) Evaluation of models for earthquake source spectra in eastern North America. *Bulletin of the Seismological Society of America* 88: 917–934.

Atkinson GM and Boore DM (2000) Reply to comment by R. A. W. Haddon on "evaluation of models for earthquake source spectra in eastern North America" by Gail M. Atkinson and David M. Boore. *Bulletin of the Seismological Society of America* 90(5): 1339–1341.

Atkinson G and Cassidy J (2000) Integrated use of seismograph and strong motion data to determine soil amplification in the Fraser delta: Results from the Duvall and Georgia Strait earthquakes. *Bulletin of the Seismological Society of America* 90: 1028–1040.

Atkinson G and Mereu R (1992) The shape of ground motion attenuation curves in southeastern Canada. *Bulletin of the Seismological Society of America* 82: 2014–2031.

Atkinson GM and Silva W (1997) An empirical study of earthquake source spectra for California earthquakes. *Bulletin of the Seismological Society of America* 87: 97–113.

Atkinson GM and Silva W (2000) Stochastic modeling of California ground motions. *Bulletin of the Seismological Society of America* 90(2): 255–274.

Atkinson G and Somerville P (1994) Calibration of time history simulation methods. *Bulletin of the Seismological Society of America* 84: 400–414.

Baltay AS, Hanks TC, and Beroza GC (2013) Stable stress-drop measurements and their variability: Implications for ground-motion prediction. *Bulletin of the Seismological Society of America* 103: 211–222. http://dx.doi.org/10.1785/0120120161.

Baltay A, Ide S, Prieto G, and Beroza G (2011) Variability in earthquake stress drop and apparent stress. *Geophysical Research Letters* 38: L06303. http://dx.doi.org/10.1029/2011GL046698.

Baltay A, Prieto G, and Beroza GC (2010) Radiated seismic energy from coda measurements and no scaling in apparent stress with seismic moment. *Journal of Geophysical Research* 115: B08314. http://dx.doi.org/10.1029/2009JB006736.

Bao HS, Bielak J, Ghattas O, et al. (1998) Large-scale simulation of elastic wave propagation in heterogeneous media on parallel computers. *Computer Methods in Applied Mechanics and Engineering* 152(1–2): 85–102.

Bard P-Y and Bouchon M (1985) The two-dimensional resonance of sediment-filled valleys. *Bulletin of the Seismological Society of America* 75: 519–541.

Bard P-Y and Tucker BE (1985) Ridge and tunnel effects: Comparing observations with theory. *Bulletin of the Seismological Society of America* 75: 905–922.

Bay F, Fäh D, Malagnini L, and Giardini D (2003) Spectral shear-wave ground-motion scaling in Switzerland. *Bulletin of the Seismological Society of America* 93: 414–429.

Beauval C, Tasan H, Laurendeau A, et al. (2012) On the testing of ground motion prediction equations against small magnitude data. *Bulletin of the Seismological Society of America* 102(5): 1994–2007.

Beeler NM, Wong T-F, and Hickman SH (2003) On the expected relationships among apparent stress, static stress drop, effective shear fracture energy, and efficiency. *Bulletin of the Seismological Society of America* 93: 1381–1389.

Bell JW, Brune JN, Liu TZ, Zreda M, and Yount JC (1998) Dating precariously balanced rocks in seismically active parts of California and Nevada. *Geology* 26(6): 495–498.

Benetatos CA and Kiratzi AA (2004) Stochastic strong ground motion simulation of intermediate depth earthquakes: The cases of the 30 May 1990 Vrancea (Romania) and of the 22 January 2002 Karpathos Island (Greece) earthquakes. *Soil Dynamics and Earthquake Engineering* 24(1): 1–9.

Benites R and Olsen KB (2005) Modeling strong ground motion in the Wellington metropolitan area, New Zealand. *Bulletin of the Seismological Society of America* 95: 2180–2196.

Ben-Zion Y (1998) Properties of seismic fault zone waves and their utility for imaging low-velocity structures. *Journal of Geophysical Research* 103: 12567–12585.

Beresnev IA and Atkinson GM (1998) Stochastic finite-fault modeling of ground motions from the 1994 Northridge, California, earthquake. I. Validation on rock sites. *Bulletin of the Seismological Society of America* 88(6): 1392–1401.

Beresnev I and Atkinson G (1999) Generic finite-fault model for ground motion prediction in eastern North America. *Bulletin of the Seismological Society of America* 89: 608–625.

Beresnev IA and Atkinson GM (2002) Source parameters of earthquakes in eastern and western North America based on finite-fault modeling. *Bulletin of the Seismological Society of America* 92(2): 695–710.

Beresnev IA, Atldnson GM, Johnson PA, and Field EH (1998) Stochastic finite-fault modeling of ground motions from the 1994 Northridge, California earthquake. II. Widespread nonlinear response at soil sites. *Bulletin of the Seismological Society of America* 88: 1402–1410.

Beresnev IA, Field EH, Van Den Abeele K, and Johnson PA (1998) Magnitude of nonlinear sediment response in Los Angeles basin during the 1994 Northridge, California, earthquake. *Bulletin of the Seismological Society of America* 88: 1079–1084.

Bernard P and Madariaga R (1984) A new asymptotic method for the modeling of near-field accelerograms. *Bulletin of the Seismological Society of America* 74: 539–557.

Beroza GC and Mikumo T (1996) Short slip duration in dynamic rupture in the presence of heterogeneous fault properties. *Journal of Geophysical Research* 101: 22449–22460.

Biasi G and Anderson JG (2007) Measurement of the parameter kappa, and reevaluation of kappa for small to moderate earthquakes at seismic stations in the vicinity of Yucca Mountain, Nevada. Technical Report TR-07-007, NSHE Qualified Reports. http://hrc.nevada.edu/qa/report.htm.

Biasi G and Smith KD (2001) Final project report, "Site effects for seismic monitoring stations in the vicinity of Yucca Mountain, Nevada", prepared for the US DOE/ UCCSN Cooperative Agreement Number DE-FC08-98NV12081, Task 12, Southern Great Basin Seismic Network Operations (SGBSN), MOL.20011204.0045.

Boatwright J (1994) Regional propagation characteristics and source parameters of earthquakes in northeastern North America. *Bulletin of the Seismological Society of America* 84(1): 1–15.

Boatwright J and Boore DM (1982) Analysis of the ground accelerations radiated by the 1980 Livermore Valley earthquakes for directivity and dynamic source characteristics. *Bulletin of the Seismological Society of America* 72(6): 1843–1865.

Boatwright J and Choy GL (1992) Acceleration source spectra anticipated for large earthquakes in northeastern North America. *Bulletin of the Seismological Society of America* 82: 660–682.

Bolt B (1969) Duration of strong motion. In: *Proceedings of the 4th World Conference on Earthquake Engineering*, Santiago, Chile, pp. 1304–1315.

Bolt BA and Abrahamson NA (2003) Estimation of strong seismic ground motions. In: Lee WHK, Kanamori H, and Jennings PC (eds.) *International Handbook of Earthquake and Engineering Seismology*, vol. 81B, pp. 983–1001. Amsterdam: International Association of Seismology and Physics of the Earth's Interior, Committee on Education, Academic Press.

Bommer JJ, Abrahamson NA, Strasser FO, et al. (2004) The challenge of defining upper bounds on earthquake ground motions. *Seismic Research Letters* 75(1): 82–95.

Bonilla LF, Archuleta RJ, and Lavallée D (2005) Hysteretic and dilatant behavior of cohesionless soils and their effects on nonlinear site response: Field data observations and modeling. *Bulletin of the Seismological Society of America* 95(6): 2373–2395. http://dx.doi.org/10.1785/0120040128.

Bonilla LF, Steidl JH, Gariel J-C, and Archuleta RJ (2002) Borehole response studies at the Garner Valley downhole array, Southern California. *Bulletin of the Seismological Society of America* 92: 3165–3179.

Boore DM (1983) Stochastic simulation of high-frequency ground motions based on seismological models of the radiated spectra. *Bulletin of the Seismological Society of America* 73: 1865–1894.

Boore DM (1986) Short-period P- and S-wave radiation from large earthquakes: Implications for spectral scaling relations. *Bulletin of the Seismological Society of America* 76(1): 43–64.

Boore DM (2010) Orientation-independent, nongeometric-mean measures of seismic intensity from two horizontal components of motion. *Bulletin of the Seismological Society of America* 100: 1830–1835.

Boore DM and Atkinson GM (1987) Stochastic prediction of ground motion and spectral response parameters at hard-rock sites in eastern North America. *Bulletin of the Seismological Society of America* 77: 440–467.

Boore D and Atkinson G (1989) Spectral scaling of the 1985–1988 Nahanni, Northwest Territories, earthquakes. *Bulletin of the Seismological Society of America* 79: 1736–1761.

Boore D and Atkinson G (1992) Source spectra for the 1988 Saguenay, Quebec earthquakes. *Bulletin of the Seismological Society of America* 82: 683–719.

Boore DM and Atkinson GM (2008) Ground-motion prediction equations for the average horizontal component of PGA, PGV, and 5%-damped PSA at spectral periods between 0.01 s and 10 s. *Earthquake Spectra* 24: 99–138.

Boore DM and Bommer JJ (2005) Processing of strong-motion accelerograms: Needs, options and consequences. *Soil Dynamics and Earthquake Engineering* 25: 93–115.

Boore DM and Joyner WB (1997) Site amplifications for generic rock sites. *Bulletin of the Seismological Society of America* 87: 327–341.

Boore DM, Stephens CD, and Joyner WB (2002) Comments on baseline correction of digital strong-motion data: Examples from the 1999 Hector Mine, California, earthquake. *Bulletin of the Seismological Society of America* 92(4): 1543–1560.

Boore DM, Watson-Lamprey J, and Abrahamson NA (2006) Orientation-independent measures of ground motion. *Bulletin of Seismological Society of America* 96: 1502–1511.

Borcherdt RD and Glassmoyer GN (1994) Influences of local geology on strong and weak ground motions recorded in the San Francisco Bay region and their implications for site-specific building-code provisions. In: Borcherdt RD (ed.) U.S. *Geological Survey Professional Paper 1551-A*, pp. A77–A108. Washington: U. S. Government Printing Office.

Borcherdt RD and Glassmoyer G (1992) On the characteristics of local geology and their influence on ground motions generated by the Loma Prieta earthquake in the San Francisco Bay region, California. *Bulletin of the Seismological Society of America* 82: 603–641.

Bouchon M and Barker JS (1996) Seismic response of a hill: the example of Tarzana, California. *Bulletin of Seismological Society of America* 86: 66–72.

Branzanti M, Colosimo G, Crespi M, et al. (2013) GPS near-real-time coseismic displacements for the great Tohoku-oki earthquake. *IEEE Geoscience and Remote Sensing Letters* 10(2): 372–376. http://dx.doi.org/10.1109/LGRS.2012.2207704.

Briggs RW, Wesnousky SG, Brune JN, Purvance MD, and Mahan SA (2013) Low footwall accelerations and variable surface rupture behavior on the Fort Sage Mountains Fault, Northeast California. *Bulletin of the Seismological Society of America* 103(1): 157–168.

Brune JN (1970) Tectonic stress and the spectra of seismic shear waves from earthquake. *Journal of Geophysical Research* 75: 4997–5009.

Brune JN (1971) Correction. *Journal of Geophysical Research* 76: 5002.

Brune JN (1976) The physics of earthquake strong motion. In: Lomnitz C and Rosenblueth E (eds.) *Seismic Risk and Engineering Decisions*, pp. 141–177. New York: Elsevier.

Brune JN (1996) Precariously balanced rocks and ground-motion maps for southern California. *Bulletin of the Seismological Society of America* 86(1): 43–54.

Brune JN (2000) Precarious rock evidence for low ground shaking on the footwall of major normal faults. *Bulletin of the Seismological Society of America* 90(4): 1107–1112.

Brune JN (2001) Shattered rock and precarious rock evidence for strong asymmetry in ground motions during thrust faulting. *Bulletin of the Seismological Society of America* 91(3): 441–447.

Brune JN (2002) Precarious-rock constraints on ground motion from historic and recent earthquakes in southern California. *Bulletin of the Seismological Society of America* 92(7): 2602–2611.

Brune JN (2003) Precarious rock evidence for low near-source accelerations for trans-tensional strike-slip earthquakes. *Physics of the Earth and Planetary Interiors* 137(1–4): 229–239.

Brune JN and Allen CR (1967) A micro-earthquake survey of the San Andreas fault system in southern California. *Bulletin of the Seismological Society of America* 57(2): 277–296.

Brune JN, Anooshehpoor A, Purvance MD, and Brune RJ (2006) Band of precariously balanced rocks between the Elsinore and San Jacinto, California, fault zones: Constraints on ground motion for large earthquakes. *Geology* 34(3): 137–140.

Brune JN, Anooshehpoor A, Shi BP, and Zeng Y (2004) Precarious rock and overturned transformer evidence for ground shaking in the M-s 7.7 Kern County earthquake: An analog for disastrous shaking from a major thrust fault in the Los Angeles basin. *Bulletin of the Seismological Society of America* 94(6): 1993–2003.

Brune JN, Smith KD, and Zeng YH (2005) Precarious rock evidence for seismic shaking during and prior to the 1992 M-L 5.6 Little Skull Mountain, Nevada, earthquake. *Earthquake Spectra* 21(4): 967–985.

Brune JN, von Seggern D, and Anooshehpoor A (2003) Distribution of precarious rocks at the Nevada Test Site: Comparison with ground motion predictions from nuclear tests. *Journal to Geophysical Research* 108(B6): 2306. http://dx.doi.org/10.1029/2002JB002000.

Building Seismic Safety Council (1997) *NEHRP Recommended Provisions for Seismic Regulations for New Buildings and Other Structures*, 1997 Edition, Part 1: Provisions (FEMA 302), 336 pp. Washington, DC: Building Seismic Safety Council.

Campbell KW (2003) Strong-motion attenuation equations. In: Lee WHK, Kanamori H, and Jennings PC (eds.) *International Handbook of Earthquake and Engineering Seismology*, vol. 81B, pp. 1003–1012. Amsterdam: International Association of Seismology and Physics of the Earth's Interior, Committee on Education, Academic Press.

Campbell KW and Bozorgnia Y (2008) NGA ground motion model for the geometric mean horizontal component of PGA, PGV, PGD and 5% damped linear elastic response spectra for periods ranging from 0.01 to 10 s. *Earthquake Spectra* 24: 137–171.

Castro RR, Anderson JG, and Singh SK (1990) Site response, attenuation and source spectra of S-waves along the Guerrero, Mexico subduction zone. *Bulletin of the Seismological Society of America* 79: 1481–1503.

Chanerley AA, Alexander NA, Berrill J, Avery H, Halldorsson B, and Sigbjornsson R (2013) Concerning baseline errors in the form of acceleration transients when recovering displacements from strong motion records using the undecimated

wavelet transform. *Bulletin of the Seismological Society of America* 103(1): 283–295.
Chapman CH (2002) Seismic ray theory and finite frequency extensions. In: Lee WHK, Kanamori H, and Jennings PC (eds.) *International Handbook of Earthquake and Engineering Seismology*, vol. 81A, pp. 103–123. Amsterdam: International Association of Seismology and Physics of the Earth's Interior, Committee on Education, Academic Press.
Chen S-Z and Atkinson GM (2002) Global comparisons of earthquake source spectra. *Bulletin of the Seismological Society of America* 92(3): 885–895.
Chi WC, Dreger D, and Kaverina A (2001) Finite-source modeling of the 1999 Taiwan (Chi-Chi) earthquake derived from a dense strong-motion network. *Bulletin of the Seismological Society of America* 91(5): 1144–1157.
Chin BH and Aki K (1991) Simultaneous study of source, path and site effects on strong ground motion during the 1989 Loma Prieta earthquake: A preliminary result on pervasive nonlinear site effects. *Bulletin of the Seismological Society of America* 81: 1859–1884.
Chiou B, Darragh R, Gregor N, and Silva W (2008) NGA project strong-motion database. *Earthquake Spectra* 24: 23–44.
Chiou BS-J and Youngs RR (2008) An NGA model for the average horizontal component of peak ground motion and response spectra. *Earthquake Spectra* 24: 173–215.
Christensen N (1989) Seismic properties of rocks. In: James DE (ed.) *The Encyclopedia of Solid Earth Geophysics*, pp. 1093–1098. New York: Van Nostrand Reinhold Company.
Cormier VF (1982) The effect of attenuation on seismic body waves. *Bulletin of the Seismological Society of America* 72: S169–S200.
Cultrera GC, Boore DM, Joyner WB, and Dietel CM (1999) Nonlinear soil response in the vicinity of the Van Norman Complex following the 1994 Northridge, California, earthquake. *Bulletin of the Seismological Society of America* 89: 1214–1231.
Dan K, Watanabe T, Tanaka T, and Sato R (1990) Stability of earthquake ground motion synthesized by using different small-event records as empirical Green's functions. *Bulletin of the Seismological Society of America* 80: 1433–1455.
Davis PM, Rubinstein JL, Liu KH, Gao SS, and Knopoff L (2000) Northridge earthquake damage caused by geologic focusing of seismic waves. *Science* 289: 1746–1750.
Delavaud E, Scherbaum F, Kuehn N, and Allen T (2012) Testing the global applicability of ground-motion prediction equations for active shallow crustal regions. *Bulletin of the Seismological Society of America* 102(2): 707–721.
Der ZA, Marshall ME, O'Donnell A, and Mcelfresh TW (1984) Spatial coherence structure and attenuation of the Lg phase, site effects, and the interpretation of the Lg coda. *Bulletin of the Seismological Society of America* 74(4): 1125–1147.
Di Alessandro C, Bonilla LF, Boore DM, Rovelli A, and Scotti O (2012) Predominant-period site classification for response spectra prediction equations in Italy. *Bulletin of the Seismological Society of America* 102(2): 680–695.
Doser D and Kanamori H (1986) Depth of seismicity in the Imperial Valley Region (1977–1983) and its relationship to heat flow, crustal structure, and the October 15, 1979 earthquake. *Journal of Geophysical Research* 91: 675–688.
Douglas J (2011) Ground-motion prediction equations 1964–2010, 455 pp. Report PEER 2011/102. Berkeley, CA: Pacific Earthquake Engineering Research Center.
Duan B and Oglesby DD (2005) The dynamics of thrust and normal faults over multiple earthquake cycles: Effects of dipping fault geometry. *Bulletin of the Seismological Society of America* 95: 1623–1636.
Durek JJ and Ekstrom G (1996) A radial model of anelasticity consistent with long-period surface-wave attenuation. *Bulletin of the Seismological Society of America* 86: 144–158.
Edwards B and Fäh D (2013) A stochastic ground-motion model for Switzerland. *Bulletin of the Seismological Society of America* 103(1): 78–98.
Elgamal AW (1991) Shear hysteretic elastoplastic earthquake response of soil systems. *Earthquake Engineering and Structural Dynamics* 20(4): 371–387.
Erickson D, McNamara DE, and Benz HM (2004) Frequency-dependent Lg Q within the continental United States. *Bulletin of the Seismological Society of America* 94: 1630–1643.
Esposito S and Iervolino I (2011) PGA and PGV spatial correlation models based on European multievent datasets. *Bulletin of the Seismological Society of America* 101(5): 2532–2541.
Esposito S and Iervolino I (2012) Spatial correlation of spectral acceleration in European data. *Bulletin of the Seismological Society of America* 102(6): 2781–2788.
Faccioli E and Pessina V (2003) Use of engineering seismology tools in ground shaking scenarios. In: Lee WHK, Kanamori H, and Jennings PC (eds.) *International Handbook of Earthquake and Engineering Seismology*, vol. 81B, pp. 1031–1048. Amsterdam: International Association of Seismology and Physics of the Earth's Interior, Committee on Education, Academic Press.
Field EH, Johnson PA, Beresnev IA, and Zeng Y (1997) Nonlinear ground-motion amplifications by sediments during the 1994 Northridge earthquake. *Nature* 390: 599–602.
Frankel A (1991) High frequency spectral falloff of earthquakes, fractal dimension of complex rupture, b-value, and the scaling of stress on faults. *Journal of Geophysical Research* 96: 6291–6302.
Frankel A (1993) Three-dimensional simulations of ground motions in the San Bernardino Valley, California, for hypothetical earthquakes on the San Andreas fault. *Bulletin of the Seismological Society of America* 83: 1020–1041.
Frankel A (1994) Dense array recordings in the San Bernardino Valley of landers-big bear aftershocks: Basin surface waves, Moho reflections, and three-dimensional simulations. *Bulletin of the Seismological Society of America* 84: 613–624.
Frankel A (1995) Simulating strong motions of large earthquakes using recordings of small earthquakes: The Loma Prieta mainshock as a test case. *Bulletin of the Seismological Society of America* 85: 1144–1160.
Frankel A, Carver D, Cranswick E, Bice T, Sell R, and Hanson S (2001) Observations of basin ground motions from a dense seismic array in San Jose, California. *Bulletin of the Seismological Society of America* 91: 1–12.
Frankel AD, Carver DL, and Williams RA (2002a) Nonlinear and linear site response and basin effects in Seattle for the M 6.8 Nisqually, Washington, earthquake. *Bulletin of the Seismological Society of America* 92: 2090–2109.
Frankel A and Clayton RW (1984) A finite-difference simulation of wave propagation in two-dimensional random media. *Bulletin of the Seismological Society of America* 74: 2167–2186.
Frankel A, Fletcher J, Vernon F, et al. (1986) Rupture characteristics and tomographic source imaging of the $M_L \sim 3$ earthquakes near Anza, California. *Journal of Geophysical Research* 91: 12633–12650.
Frankel A and Stephenson W (2000) Three-dimensional simulations of ground motions in the Seattle region for earthquakes in the Seattle Fault zone. *Bulletin of the Seismological Society of America* 90: 1251–1267.
Frankel A and Vidale J (1992) A 3-dimensional simulation of seismic-waves in the Santa-Clara Valley, California, from a Loma-Prieta aftershock. *Bulletin of the Seismological Society of America* 82(5): 2045–2074.
Frankel A and Wennerberg L (1987) Energy-flux model of seismic coda: Separation of scattering and intrinsic attenuation. *Bulletin of the Seismological Society of America* 77: 1223–1251.
Frankel AD, Petersen MD, Mueller CS, et al. (2002b) Documentation for the 2002 update of the national seismic hazard maps. Open File Report 02-420. Denver, CO: US Geological Survey.
Furumura T and Chen L (2004) Large scale parallel simulation and visualization of 3D seismic wavefield using the Earth Simulator. *Computer Modeling in Engineering and Sciences* 6: 149–165.
Furumura T and Hayakawa T (2007) Anomalous propagation of long-period ground motions recorded in Tokyo during the 23 October 2004 Mw 6.6 Niigata-ken Chuetsu, Japan, earthquake. *Bulletin of the Seismological Society of America* 97: 863–880. http://dx.doi.org/10.1785/0120060166.
Furumura T, Kennett BLN, and Koketsu K (2003) Visualization of 3D wave propagation from the 2000 Tottori-ken Seibu, Japan, Earthquake: Observation and numerical simulation. *Bulletin of the Seismological Society of America* 93(2): 870–881. [Online movies are available for the 23 October, 2004, Chuetsu earthquake (M=6.8) at (http://www.eri.u-tokyo.ac.jp/furumura/04Chuetsu/index.html) and for the 16 August 2005 Miyagioki, Japan, earthquake (M=7.2) at (http://www.eri.u-tokyo.ac.jp/furumura/05Miyagi/).
Furumura T and Koketsu K (1998) Specific distribution of ground motion during the 1995 Kobe earthquake and its generation mechanism. *Geophysical Research Letters* 25: 785–788.
Furumura T and Singh SK (2002) Regional wave propagation from Mexican subduction zone earthquakes: The attenuation functions for interplate and inslab events. *Bulletin of the Seismological Society of America* 92: 2110–2125.
Gao S, Liu H, Davis PM, and Knopoff L (1996) Localized amplification of seismic waves and correlation with damage due to the Northridge earthquake: Evidence for focusing in Santa Monica. *Bulletin of the Seismological Society of America* 86: S209–S230.
Gao Y, Zhang N, Li D, Liu H, Cai Y, and Wu Y (2012) Effects of topographic amplification induced by a U-shaped canyon on seismic waves. *Bulletin of the Seismological Society of America* 102(4): 1748–1763.
García D, Singh SK, Herráiz M, Ordaz M, and Pacheco JF (2005) Inslab earthquakes of Central Mexico: Peak ground-motion parameters and response spectra. *Bulletin of the Seismological Society of America* 95: 2272–2282.
Geli L, Bard PY, and Julien B (1988) The effect of topography on earthquake ground motion: A review and new results. *Bulletin of the Seismological Society of America* 78: 42–63.

Genrich JF and Bock Y (2006) Instantaneous geodetic positioning with 10–50 Hz GPS measurements: Noise characteristics and implications for monitoring networks. *Journal of Geophysical Research* 111: B03403. http://dx.doi.org/10.1029/2005JB003617.

Gil-Zepeda SA, Luzón F, Aguirre J, Morales J, Sánchez-Sesma FJ, and Ortiz-Alemán C (2002) 3D seismic response of the deep basement structure of the Granada Basin (southern Spain). *Bulletin of the Seismological Society of America* 92: 2163–2176.

Goda K and Hong HP (2008) Spatial correlation of peak ground motions and response spectra. *Bulletin of the Seismological Society of America* 98(1): 354–365.

Goulet C (2013) Development and implementation of validation exercises using the SCEC broadband strong ground motion simulation platform (abstract). *Seismological Research Letters* 84(2): 336.

Graves RW (1993) Modeling three-dimensional site response effects in the Marina district basin, San Francisco, California. *Bulletin of the Seismological Society of America* 83: 1042–1063.

Graves RW (1996) Simulating seismic wave propagation in 3D elastic media using staggered-grid finite differences. *Bulletin of the Seismological Society of America* 86(4): 1091–1106.

Graves RW (1998) Three-dimensional finite-difference modeling of the San Andreas fault: Source parameterization and ground-motion levels. *Bulletin of the Seismological Society of America* 88: 881–897.

Graves RW, Pitarka A, and Somerville PG (1998) Ground motion amplification in the Santa Monica area: Effects of shallow basin edge structures. *Bulletin of the Seismological Society of America* 88: 1224–1242.

Graves RW and Wald DJ (2004) Observed and simulated ground motions in the San Bernardino basin region for the Hector Mine, California, earthquake. *Bulletin of the Seismological Society of America* 94: 131–146.

Gregor NJ, Silva WJ, Wong IG, and Youngs RR (2002) Ground-motion attenuation relationships for Cascadia subduction zone megathrust earthquakes based on a stochastic finite-fault model. *Bulletin of the Seismological Society of America* 92(5): 1923–1932.

Guatteri M, Martin Mai P, and Beroza GC (2004) A pseudo-dynamic approximation to dynamic rupture models for strong ground motion prediction. *Bulletin of the Seismological Society of America* 94: 2051–2063.

Guatteri M, Martin Mai P, Beroza GC, and Boatwright J (2003) Strong ground-motion prediction from stochastic-dynamic source models. *Bulletin of the Seismological Society of America* 93: 301–313.

Guguen P, Bard P-Y, and Chávez-García FJ (2002) Site-city seismic interaction in Mexico City-like environments: An analytical study. *Bulletin of the Seismological Society of America* 92(2): 794–811.

Guéguen P, Bard P-Y, and Oliveira CS (2000) Experimental and numerical analysis of soil motions caused by free vibrations of a building model. *Bulletin of the Seismological Society of America* 90(6): 1464–1479.

Haddon R (1996) Earthquake source spectra in eastern North America. *Bulletin of the Seismological Society of America* 86: 1300–1313.

Haddon R (1997) Reply to Comment on "Earthquake source spectra in eastern North America" by Atkinson, Boore, and Boatwright. *Bulletin of the Seismological Society of America* 87: 1703–1708.

Hall JF (2003) Finite element analysis in earthquake engineering. In: Lee WHK, Kanamori H, and Jennings PC (eds.) *International Handbook of Earthquake and Engineering Seismology*, vol. 81B, pp. 1133–1158. Amsterdam: International Association of Seismology and Physics of the Earth's Interior, Committee on Education, Academic Press.

Hall JF, Heaton TH, Halling MW, and Wald DJ (1995) Near-source ground motion and its effects on flexible buildings. *Earthquake Spectra* 11(4): 569–605.

Hanks TC (1979) b-values and $\omega^{-\gamma}$ seismic source models: implications for tectonic stress variations along active crustal fault zones and the estimation of high-frequency strong ground motion. *Journal of Geophysical Research* 84: 2235–2242.

Hanks TC (1982) fmax. *Bulletin of the Seismological Society of America* 72: 1867–1879.

Hanks TH and Bakun WH (2002) A bilinear source-scaling model for M-log A observations of continental earthquakes. *Bulletin of the Seismological Society of America* 92: 1841–1846.

Hanks TC and Johnson DA (1976) Geophysical assessment of peak accelerations. *Bulletin of the Seismological Society of America* 66(3): 959–968.

Hanks TC and McGuire RK (1981) The character of high-frequency strong ground motion. *Bulletin of the Seismological Society of America* 71(6): 2071–2095.

Hartzell SH (1978) Earthquake aftershocks as Green's functions. *Geophysical Research Letters* 5: 1–4.

Hartzell S, Bonilla LF, and Williams RA (2004) Prediction of nonlinear soil effects. *Bulletin of the Seismological Society of America* 94(5): 1609–1629.

Hartzell S, Carver D, Cranswick E, and Frankel A (2000) Variability of site response in Seattle, Washington. *Bulletin of the Seismological Society of America* 90: 1237–1250.

Hartzell SH, Carver DL, and King KW (1994) Initial investigation of site and topographic effects at Robinwood Ridge, California. *Bulletin of the Seismological Society of America* 84: 1336–1349.

Hartzell S, Cranswick E, Frankel A, Carver D, and Meremonte M (1997) Variability of site response in the Los Angeles urban area. *Bulletin of the Seismological Society of America* 87: 1377–1400.

Hartzell S, Harmsen S, Frankel A, and Larsen S (1999) Calculation of broadband time histories of ground motion: Comparison of methods and validation using strong-ground motion from the 1994 Northridge earthquake. *Bulletin of the Seismological Society of America* 89(6): 1484–1504.

Hartzell S, Leeds A, Frankel A, et al. (2002) Simulation of broadband ground motion including nonlinear soil effects for a magnitude 6.5 earthquake on the Seattle Fault, Seattle, Washington. *Bulletin of the Seismological Society of America* 92: 831–853.

Hartzell S, Mariagiovanna G, Mai PM, Liu P-C, and Fisk M (2005) Calculation of broadband time histories of ground motion, Part II: Kinematic and dynamic modeling using theoretical Green's functions and comparison with the 1994 Northridge earthquake. *Bulletin of the Seismological Society of America* 95: 614–645.

Haskell NA (1953) The dispersion of surface waves on multilayered media. *Bulletin of the Seismological Society of America* 43: 17–34.

Haskell NA (1960) Crustal reflection of plane SH waves. *Journal of Geophysical Research* 65: 4147–4150.

Haskell NA (1962) Crustal reflection of plane P and SV waves. *Journal of Geophysical Research* 67: 4751–4767.

Haskell NA (1964) Total energy and energy spectral density of elastic wave radiation from propagating faults. *Bulletin of the Seismological Society of America* 54: 1811–1841.

Haskell NA (1969) Elastic displacements in the near-field of a propagating fault. *Bulletin of the Seismological Society of America* 59: 865–908.

Hayashi K, Burns DR, and Toksoz MN (2001) Discontinuous-grid finite-difference seismic modeling including surface topography. *Bulletin of the Seismological Society of America* 91(6): 1750–1764.

Heaton TH (1990) Evidence for and implications of self-healing pulses of slip in earthquake rupture. *Physics of the Earth and Planetary Interiors* 64: 1–20.

Heaton TH, Hall JF, Wald DJ, and Halling MJ (1995) Response of high-rise and base-isolated buildings to a hypothetical M (sub W) 7.0 blind thrust earthquake. *Science* 267(5195): 206–211.

Higashi S and Sasatani T (2000) Nonlinear site response in Kushiro during the 1994 Hokkaido Toho-oki earthquake. *Bulletin of the Seismological Society of America* 90(4): 1082–1095.

Hisada Y and Bielak J (2003) A theoretical method for computing near-fault ground motions in layered half-spaces considering static offset due to surface faulting, with a physical interpretation of fling step and rupture directivity. *Bulletin of the Seismological Society of America* 93(3): 1154–1168.

Horike M (1985) Inversion of phase velocity of long-period microtremors to the S-wave-velocity structure down to the basement in urbanized areas. *Journal of Physics of the Earth* 33: 59–96.

Hough S (1996) Observational constraints on earthquake source scaling: Understanding the limits in resolution. *Tectonophysics* 261: 83–95.

Hough SE, Altidor JR, Anglade D, et al. (2010) Localized damage caused by topographic amplification during the 2010 M7.0 Haiti earthquake. *Nature Geoscience* 3: 778–782.

Hough SE and Anderson JG (1988) High-frequency spectra observed at Anza, California: Implications for Q structure. *Bulletin of the Seismological Society of America* 78: 692–707.

Hough SE and Field EH (1996) On the coherence of ground motion in the San Fernando Valley. *Bulletin of the Seismological Society of America* 86(6): 1724–1732.

Hough SE, Friberg PA, Busby R, Field EF, Jacob KH, and Borcherdt RD (1990) Sediment-induced amplification and the collapse of the Nimitz Freeway. *Nature* 344: 853–855.

Hough SE, Seeber L, Rovelli A, et al. (1992) Ambient noise and weak-motion excitation of sediment resonances: Results from the Tiber Valley, Italy. *Bulletin of the Seismological Society of America* 82: 1186–1205.

Hudson DE (1963) The measurement of ground motion of destructive earthquakes. *Bulletin of the Seismological Society of America* 53: 419–437.

Hudson DE (1979) *Reading and Interpreting Strong Motion Accelerograms*. Berkeley, CA: Earthquake Engineering Research Institute, 112 pp.

Humphrey JR Jr. and Anderson JG (1992) Shear wave attenuation and site response in Guerrero, Mexico. *Bulletin of the Seismological Society of America* 82: 1622–1645.

Humphrey J and Anderson J (1994) Seismic source parameters from the Guerrero subduction zone. *Bulletin of the Seismological Society of America* 84(6): 1754–1769.

Hutchings L (1991) "Prediction" of strong ground motion for the 1989 Loma Prieta earthquake using empirical Green's functions. *Bulletin of the Seismological Society of America* 81(5): 1813–1837.

Hutchings L (1994) Kinematic earthquake models and synthesized ground motion using empirical Green's functions. *Bulletin of the Seismological Society of America* 84: 1028–1050.

Hwang R-D, Wang J-H, Huang B-S, et al. (2001) Estimates of stress drop of the Chi-Chi, Taiwan, earthquake of 20 September 1999 from near-field seismograms. *Bulletin of the Seismological Society of America* 91(5): 1158–1166.

Ide S, Baltay A, and Beroza GC (2011) Shallow dynamic overshoot and energetic deep rupture in the 2011 M_W9.0 Tohoku-Oki earthquake. *Science* 332: 1426–1429.

Ide S and Beroza GC (2001) Does apparent stress vary with earthquake size? *Geophysical Research Letters* 28(17): 3349–3352. http://dx.doi.org/10.1029/2001GL013106.

Idriss IM (2008) An NGA empirical model for estimating the horizontal spectral values generated by shallow crustal earthquakes. *Earthquake Spectra* 24: 217–242.

Inoue T and Miyatake T (1997) 3-D simulation of near-field strong ground motion: Basin edge effect derived from rupture directivity. *Geophysical Research Letters* 24: 905–908.

Irikura K (1983) Semi-empirical estimation of strong ground motions during large earthquakes. *Bulletin of the Disaster Prevention Research Institute* 33(2): 63–104.

Iwaki A and Iwata T (2011) Estimation of three-dimensional boundary shape of the Osaka sedimentary basin by waveform inversion. *Geophysical Journal International* 186: 1255–1278. http://dx.doi.org/10.1111/j.1364-246X.2011.05102.x.

Iwan WI (1979) The deployment of strong-motion earthquake instrument arrays. *Earthquake Engineering and Structural Dynamics* 7: 413–426.

Jaeger JC and Cook NGW (1976) *Fundamentals of Rock Mechanics*, 2nd edn. London: Chapman and Hall, 585 pp.

Javelaud EH, Ohmachi T, and Inoue S (2011) A quantitative approach for estimating coseismic displacements in the near field from strong motion accelerographs. *Bulletin of the Seismological Society of America* 101: 1182–1198.

Jayaram N and Baker JW (2009) Correlation model for spatially distributed ground-motion intensities. *Earthquake Engineering and Structural Dynamics* 38(15): 1687–1708.

Jennings PC (1970) Distant motions from a building vibration test. *Bulletin of the Seismological Society of America* 60(6): 2037–2043.

Jennings PC (2003) An introduction to the earthquake response of structures. *International Handbook of Earthquake and Engineering Seismology*, vol. 81B, pp. 1097–1125. London: International Association for Seismology and Physics of the Earth's Interior, Academic Press.

Jibson RW and Harp EL (2012) Extraordinary distance limits of landslides triggered by the 2011 Mineral, Virginia, earthquake. *Bulletin of the Seismological Society of America* 102: 2368–2377.

Jones LE and Helmberger DV (1998) Earthquake source parameters and fault kinematics in the eastern California shear zone. *Bulletin of the Seismological Society of America* 88(6): 1337–1352.

Joswig M (2001) Mapping seismic quiescence in California. *Bulletin of the Seismological Society of America* 91(1): 64–81.

Joyner WB (1975) A method for calculating nonlinear seismic response in two-dimensions. *Bulletin of the Seismological Society of America* 65: 1337–1357.

Joyner WB (1984) A scaling law for the spectra of large earthquakes. *Bulletin of the Seismological Society of America* 74: 1167–1188.

Joyner WB and Boore DM (1986) On simulating large earthquakes by Green's-function addition of smaller earthquakes. In: Das S, Boatwright J, and Scholz CH (eds.) *Earthquake Source Mechanics*. *Geophysical Monograph 37*pp. 269–274. Washington, DC: American Geophysical Union.

Joyner WB and Chen TF (1975) Calculation of nonlinear ground response in earthquakes. *Bulletin of the Seismological Society of America* 65: 1315–1336.

Junn J-G, Jo N-D, and Baag C-E (2002) Stochastic prediction of ground motions in southern Korea. *Geosciences Journal (Seoul)* 6(3): 203–214.

Kamae K, Irikura K, and Pitarka A (1998) A technique for simulating strong ground motion using hybrid Green's function. *Bulletin of the Seismological Society of America* 88: 357–367.

Kanamori H (1977) The energy release in great earthquakes. *Journal of Geophysical Research* 82: 2981–2987.

Kanamori H and Allen CR (1986) Earthquake repeat time and average stress drop. In: Das S, Boatwright J, and Scholz CH (eds.) *Earthquake Source Mechanics*. *Geophysical Monograph 37,* pp. 227–235. Washington, DC: American Geophysical Union.

Kanamori H and Anderson DL (1975) Theoretical basis of some empirical relations in seismology. *Bulletin of the Seismological Society of America* 65: 1073–1095.

Kanamori H, Jennings PC, Singh SK, and Astiz L (1993) Estimation of strong ground motions in Mexico City expected for large earthquakes in the Guerrero seismic gap. *Bulletin of the Seismological Society of America* 83(3): 811–829.

Kanamori H and Rivera L (2004) Static and dynamic scaling relations for earthquake and their implications for rupture speed and stress drop. *Bulletin of the Seismological Society of America* 94: 314–319.

Kawase H (1996) The cause of the damage belt in Kobe: "The basin-edge effect," constructive interference of the direct S-wave with the basin-induced diffracted/Rayleigh waves. *Seismological Research Letters* 67(5): 25–34.

Kawase H (2003) Site effects on strong ground motions. In: Lee WHK, Kanamori H, and Jennings PC (eds.) *International Handbook of Earthquake and Engineering Seismology*, vol. 81B, pp. 1013–1030. London: International Association of Seismology and Physics of the Earth's Interior, Committee on Education, Academic Press.

Kawase H, Matsushima S, Graves RW, and Somerville PG (1998a) Three-dimensional wave propagation analysis of simple two-dimensional basin structures with special reference to "the basin-edge effect" *Zishin (Journal of Seismological Society of Japan)* 50: 431–449 (in Japanese).

Kawase H, Sanchez-Sesma FJ, and Matsushima S (2011) The optimal use of horizontal-to-vertical spectral ratios of earthquake motions for velocity inversions based on diffuse-field theory for plane waves. *Bulletin of the Seismological Society of America* 101: 2001–2014. http://dx.doi.org/10.1785/0120100263.

Kawase H, Satoh T, Iwata T, and Irikura K (1998b) S-wave velocity structures in the San Fernando and Santa Monica areas. In: *Proceedings of the 2nd International Symposium on Effects of Surface Geology on Seismic Motions*, Tokyo, Japan, vol. 2, pp. 733–740.

Kawasumi H (1968) Introduction. In: Kawasumi H (Editor-in-Chief) *General Report on the Niigata Earthquake of 1964*, pp. 1–6. Tokyo, Japan: Tokyo Electrical Engineering College Press. http://www.iris.edu/seismo/quakes/1964niigata/.

Kham M, Semblat J-F, Bard P-Y, and Dangla P (2006) Seismic site–city interaction: Main governing phenomena through simplified numerical models. *Bulletin of the Seismological Society of America* 96(5): 1934–1951.

Kilb D, Biasi G, Anderson J, Brune J, Peng Z, and Vernon FL (2012) A comparison of spectral parameter kappa from small and moderate earthquakes using Southern California ANZA seismic network data. *Bulletin of the Seismological Society of America* 102: 284–300.

Kinoshita S (2003) Kyoshin net (K_NET), Japan. In: Lee WHK, Kanamori H, and Jennings PC (eds.) *International Handbook of Earthquake and Engineering Seismology*, vol. 81B, pp. 1049–1056. London: International Association of Seismology and Physics of the Earth's Interior, Committee on Education, Academic Press.

Kinoshita S and Ohike M (2002) Scaling relations of earthquakes that occurred in the upper part of the Philippine Sea plate beneath the Kanto region, Japan, estimated by means of borehole recordings. *Bulletin of the Seismological Society of America* 92(2): 611–624.

Kinoshita S and Takagishi M (2011) Generation and propagation of static displacement estimated using KiK-net recordings. *Earth, Planets and Space* 63: 779–783.

Komatitsch D, Liu Q, Tromp J, Suss P, Stidham C, and Shaw JH (2004) Simulations of ground motion in the Los Angeles Basin based upon the spectral-element method. *Bulletin of the Seismological Society of America* 94(1): 187–206.

Komatitsch D and Tromp J (1999) Introduction to the spectral element method for three-dimensional seismic wave propagation. *Geophysical Journal International* 139(3): 806–822.

Kramer SL (1996) *Geotechnical Earthquake Engineering*. Upper Saddle River, NJ: Prentice Hall, 653 pp.

Ktenidou O-J, Cotton F, Abrahamson N, and Anderson JG (2014) Taxonomy of kappa: A review of definitions and estimation approaches targeted to applications. *Seismological Research Letters* 85: 1–14. http://dx.doi.org/10.1785/0220130027.

Kudo K, Uetake T, and Kanno T (2000) Re-evaluation of nonlinear site response during the 1964 Niigata earthquake using the strong motion records from at Kawagishi-cho, Niigata City. In: *Proceedings of the 12th World Conference on Earthquake Engineering*, Paper 0969.

Kurihashi S and Irikura K (2011) Source model for generating strong ground motions during the 2011 off the Pacific Coast of Tohoku earthquake. *Earth, Planets and Space* 63: 571–576 (Special issue).

Larsen S (2002) Las Vegas basin seismic response project: 3-D finite-difference ground motion simulations. *Eos Transactions American Geophysical Union* 83(47), Fall Meeting Supplement, Abstract S12B-1188.

Larson KM, Bodin P, and Gomberg J (2003) Using 1-Hz GPS data to measure deformations caused by the Denali fault earthquake. *Science* 300: 1421–1424.

Lay T (2002) The Earth's interior. In: Lee WHK, Kanamori H, and Jennings PC (eds.) *International Handbook of Earthquake and Engineering Seismology*, vol. 81A, pp. 829–860. London: International Association of Seismology and Physics of the Earth's Interior, Committee on Education, Academic Press.

Lay T, Kanamori H, Ammon CJ, et al. (2012) Depth-varying rupture properties of subduction zone megathrust faults. *Journal of Geophysical Research* 117: B04311. http://dx.doi.org/10.1029/2011JB009133.

Lee WHK, Kanamori H and Jennings PC (eds.) (2003) *International Handbook of Earthquake and Engineering Seismology*, vol. 81B. Amsterdam: International Association of Seismology and Physics of the Earth's Interior, Committee on Education, Academic Press.

Leonard M (2010) Earthquake fault scaling: Self-consistent relating of rupture length, width, average displacement, and moment release. *Bulletin of the Seismological Society of America* 100: 1971–1988. http://dx.doi.org/10.1785/0120090189.

Leonard M (2012) Erratum to Earthquake fault scaling: Self-consistent relating of rupture length, width, average displacement, and moment release. *Bulletin of the Seismological Society of America* 102: 2797. http://dx.doi.org/10.1785/0120120249.

Levander AR (1988) 4th-order finite-difference P-SV seismograms. *Geophysics* 53(11): 1425–1436.

Li Y-G and Leary P (1990) Fault zone trapped seismic waves. *Bulletin of the Seismological Society of America* 80: 1245–1271.

Li Y-G, Leary P, Aki K, and Malin P (1990) Seismic trapped modes in the Oroville and San Andreas fault zones. *Science* 249: 763–766.

Liu HL and Helmberger DV (1985) The 23:19 aftershock of the 15 October 1979 Imperial Valley earthquake: More evidence for an asperity. *Bulletin of the Seismological Society of America* 75: 689–708.

Louie JN (2001) Faster, better: Shear-wave velocity to 100 meters depth from refraction microtremor arrays. *Bulletin of the Seismological Society of America* 91: 347–364.

Luco JE and Anderson JG (1983) Steady state response of an elastic half-space to a moving dislocation of finite width. *Bulletin of the Seismological Society of America* 73: 1–22.

Luco JE and Apsel RJ (1983) On the Green's functions for a layered half-space: Part I. *Bulletin of the Seismological Society of America* 73(4): 909–929.

Ma K-F, Mori J, Lee SJ, and Yu SB (2001) Spatial and temporal distribution of slip for the 1999 Chi-Chi, Taiwan, earthquake. *Bulletin of the Seismological Society of America* 91(5): 1069–1087.

Madariaga R and Olsen KB (2002) Earthquake dynamics. In: Lee WHK, Kanamori H, and Jennings PC (eds.) *International Handbook of Earthquake and Engineering Seismology*, vol. 81A, pp. 175–194. London: International Association Seismology and Physics of the Earth's Interior, Committee on Education, Academic Press.

Mai PM and Beroza GC (2000) Source scaling properties from finite-fault-rupture models. *Bulletin of the Seismological Society of America* 90: 604–615.

Mai P, Jonsson S, and Bethmann F (2006) SRCMOD – An online resource for finite-source rupture models of past earthquakes. In: *Abstracts of the Centennial Meeting of the Seismological Society of America*, 18–22 April 2006, San Francisco, CA, Abstract No. 295.

Malagnini L (1996) Velocity and attenuation structure of very shallow soils: Evidence for a frequency-dependent Q. *Bulletin of the Seismological Society of America* 86(5): 1471–1486.

Malagnini L, Akinci A, Herrmann RB, Pino NA, and Scognamiglio L (2002) Characteristics of the ground motion in Northeastern Italy. *Bulletin of the Seismological Society of America* 92: 2186–2204.

Malagnini L and Herrmann RB (2000) Ground-motion scaling in the region of the 1997 Umbria-Marche earthquake (Italy). *Bulletin of the Seismological Society of America* 90(4): 1041–1051.

Malagnini L, Herrmann RB, and Di Bona M (2000) Ground-motion scaling in the Apennines (Italy). *Bulletin of the Seismological Society of America* 90: 1062–1081.

Malagnini L, Herrmann RB, and Koch K (2000) Regional ground-motion scaling in Central Europe. *Bulletin of the Seismological Society of America* 90: 1052–1061.

Malagnini L, Herrmann RB, Mercuri A, Opice S, Biella G, and de Franco R (1997) Shear-wave velocity structure of sediments from the inversion of explosion-induced Rayleigh waves: Comparison with cross-hole measurements. *Bulletin of the Seismological Society of America* 87(6): 1413–1421.

Malagnini L, Mayeda K, Akinci A, and Bragato PL (2004) Estimating absolute site effects. *Bulletin of the Seismological Society of America* 94: 1343–1352.

Margaris BN and Boore DM (1998) Determination of and 0 from response spectra of large earthquakes in Greece. *Bulletin of the Seismological Society of America* 88: 170–182.

Matsushima T and Okada H (1990) Determination of deep geological structures under urban areas using long-period microtremors. *ButsuriTansa* 43: 21–33.

McGarr A and Fletcher JB (2003) Maximum slip in earthquake fault zones, apparent stress, and stick–slip friction. *Bulletin of the Seismological Society of America* 93: 2355–2362.

McGarr A, Green RWE, and Spottiswoode SM (1981) Strong ground motion of mine tremors: Some implications for near-source ground motion parameters. *Bulletin of the Seismological Society of America* 71: 295–319.

McGuire RK and Hanks TC (1980) RMS accelerations and spectral amplitudes of strong ground motion during the San Fernando, California earthquake. *Bulletin of the Seismological Society of America* 70(5): 1907–1919.

Mena B, Durukal E, and Erdik M (2006) Effectiveness of hybrid Green's function method in the simulation of near-field strong motion: An application to the 2004 Parkfield earthquake. *Bulletin of the Seismological Society of America* 96(4B): S183–S205. http://dx.doi.org/10.1785/0120050814.

Menglin L, Huaifeng W, Xi C, and Yongmei Z (2011) Structure-soil-structure interaction: Literature review. *Soil Dynamics and Earthquake Engineering* 31: 1724–1731.

Menke W and Chen R (1984) Numerical studies of the coda falloff rate of multiply scattered waves in randomly layered media. *Bulletin of the Seismological Society of America* 74: 1605–1614.

Menke W, Lerner-Lam AL, Dubendorff B, and Pacheco J (1990) Polarization and coherence of 5 to 30 Hz seismic wave fields at a hard-rock site and their relevance to velocity heterogeneities in the crust. *Bulletin of the Seismological Society of America* 80(2): 430–449.

Meremonte M, Frankel A, Cranswick E, Carver D, and Worley D (1996) Urban seismology – Northridge aftershocks recorded by multi-scale arrays of portable digital seismographs. *Bulletin of the Seismological Society of America* 86: 1350–1363.

Midorikawa S (1994) Case histories of upthrow of objects during earthquakes (in Japanese). *Jishin (Journal of the Seismological Society of Japan)* 47: 333–340.

Miyake H, Iwata T, and Irikura K (2003) Source characterization for broadband ground-motion simulation: Kinematic heterogeneous source model and strong motion generation area. *Bulletin of the Seismological Society of America* 93(6): 2531–2545.

Miyatake T (2000) Computer simulation of strong ground motion near a fault using dynamic fault rupture modeling: Spatial distribution of the peak ground velocity vectors. *Pure and Applied Geophysics* 157: 2063–2081.

Mogi H, Shrestha SM, Kawakami H, and Kawamura J (2013) Nonlinear soil behavior examined on the basis of propagation time observed at the KiK-net Ichinoseki-Nishi vertical array. *Bulletin of the Seismological Society of America* 103(1): 180–195.

Mori J (1996) Rupture directivity and slip distribution of the M 4.3 foreshock to the 1992 Joshua Tree earthquake, Southern California. *Bulletin of the Seismological Society of America* 86(3): 805–810.

Mori J and Frankel A (1990) Source parameters of aftershocks of the 1986 North Palm Springs earthquake. *Bulletin of the Seismological Society of America* 80: 278–295.

Nakamura T, Takenaka H, Okamoto T, and Kaneda Y (2012) FDM simulation of seismic-wave propagation for an aftershock of the 2009 Suruga Bay earthquake: Effects of ocean-bottom topography and seawater layer. *Bulletin of the Seismological Society of America* 102(6): 2420–2435.

Nazarian S and Stokoe II KH (1984) In situ shear wave velocities from spectral analysis of surface waves. In: *Proceedings of the World Conference on Earthquake Engineering*, vol. 8, San Francisco, 21–28 July.

Nielsen SB and Olsen KB (2000) Constraints on stress and friction from dynamic rupture models of the 1994 Northridge, California, earthquake. *Pure and Applied Geophysics* 157: 2029–2046.

Oglesby DD and Day SM (2002) Stochastic fault stress: Implications for fault dynamics and ground motion. *Bulletin of the Seismological Society of America* 92: 3006–3021.

Oglesby DD, Day SM, Li Y-G, and Vidale JE (2003) The 1999 Hector Mine earthquake: The dynamics of a branched fault system. *Bulletin of the Seismological Society of America* 93: 2459–2476.

Oglesby DD, Dreger DS, Harris RA, Ratchkovski N, and Hansen R (2004) Inverse kinematic and forward dynamic models of the 2002 Denali Fault Earthquake, Alaska. *Bulletin of the Seismological Society of America* 94: S214–S233.

Ohmachi T and Midorikawa S (1992) Ground-motion intensity inferred from upthrow of boulders during the 1984 Western Nagano Prefecture, Japan, earthquake. *Bulletin of the Seismological Society of America* 82: 44–60.

Ohminato T and Chouet BA (1997) A free surface boundary condition for including 3D topography in the finite-difference method. *Bulletin of the Seismological Society of America* 87: 494–515.

Oldham RD (1899) *Report on the Great Earthquake of 12th June 1897*. Calcutta: Geological Survey of India, vol. 29, 379 pp. Memoirs.

Olsen KB (2000) Site amplification in the Los Angeles Basin from three-dimensional modeling of ground motion. *Bulletin of the Seismological Society of America* 90: S77–S94.

Olsen KB, Akinci A, Rovelli A, Marra F, and Malagnini L (2006) 3D ground motion estimation in Rome, Italy. *Bulletin of the Seismological Society of America* 96: 133–146.

Olsen KB and Archuleta RJ (1996) Three-dimensional simulation of earthquakes on the Los Angeles fault system. *Bulletin of the Seismological Society of America* 86: 575–596.

Olsen KB, Archuleta RJ, and Matarese JR (1995a) Three-dimensional simulation of a magnitude 7.75 earthquake on the San Andreas fault. *Science* 270: 1628–1632.

Olsen KB, Day SM, and Bradley CR (2003) Estimation of Q for long-period (>2 sec) waves in the Los Angeles Basin. *Bulletin of the Seismological Society of America* 93: 627–638.

Olsen KB, Day SM, Minster JB, et al. (2006) Strong shaking in Los Angeles expected from southern San Andreas earthquake. *Geophysical Research Letters* 33: L07305. http://dx.doi.org/10.1029/2005GL025472.

Olsen KB, Day SM, Minster JB, Cui Y, Chourasia A, Okaya D, et al. (2008) TeraShake2: Spontaneous rupture simulations of MW7.7 earthquakes on the southern San Andreas fault. *Bulletin of Seismological Society of America* 98: 1162–1185.

Olsen KB and Mayhew JE (2010) Goodness-of-fit criteria for broadband synthetic seismograms, with application to the 2008 M 5.4 Chino Hills, California, earthquake. *Seismological Research Letters* 81: 715–723. http://dx.doi.org/10.1785/gssrl.81.5.715.

Olsen KB, Pechmann JC, and Schuster GT (1995b) Simulation of 3D elastic wave propagation in the Salt Lake basin. *Bulletin of the Seismological Society of America* 85: 1688–1710.

Ordaz M, Arboleda J, and Singh SK (1995) A scheme of random summation of an empirical Green's function to estimate ground motions from future large earthquakes. *Bulletin of the Seismological Society of America* 85(6): 1635–1647.

Pavlenko O and Irikura K (2002) Changes in shear moduli of liquefied and nonliquefied soils during the 1995 Kobe earthquake and its aftershocks at three vertical-array sites. *Bulletin of the Seismological Society of America* 92(5): 1952–1969.

Pavlenko O and Loh C-H (2005) Nonlinear identification of the soil response at Dahan downhole array site during the 1999 Chi–Chi earthquake. *Soil Dynamics and Earthquake Engineering* 25: 241–250.

Petersen MD, Frankel AD, Harmsen SC, et al. (2008) Documentation for the 2008 update of the United States national seismic hazard maps. *US Geological Survey Open-File Report 2008-1128*, 61 p.

Peyrat S and Olsen KB (2004) Nonlinear dynamic rupture inversion of the 2000 western Tottori, Japan, earthquake. *Geophysical Research Letters* 31: L05604.

Phillips WS and Aki K (1986) Site amplification of coda waves from local earthquake in central California. *Bulletin of the Seismological Society of America* 76: 627–648.

Pitarka A, Graves R, and Somerville P (2004) Validation of a 3D velocity model of the Puget Sound region based on modeling ground motion from the 28 February 2001 Nisqually earthquake. *Bulletin of the Seismological Society of America* 94: 1670–1689.

Pitarka A and Irikura K (1996) Basin structure effects on long-period strong motions in the San Fernando Valley and the Los Angeles Basin from the 1994 Northridge earthquake and an aftershock. *Bulletin of the Seismological Society of America* 86: S126–S137.

Pitarka A, Irikura K, and Iwata T (1997) Modelling of ground motion in the Higashinada (Kobe) area for an aftershock of the 1995 January 17 Hyogo-ken Nanbu, Japan, earthquake. *Geophysical Journal International* 131(2): 231–239.

Pitarka A, Irikura K, Iwata T, and Sekiguchi H (1998) Three dimensional simulation of the near fault ground motion for the 1995 Hyogoken Nanbu (Kobe), Japan, earthquake. *Bulletin of the Seismological Society of America* 88: 428–440.

Pitarka A, Somerville PG, Fukushima Y, and Uetake T (2002) Ground-motion attenuation from the 1995 Kobe earthquake based on simulations using the hybrid Green's function method. *Bulletin of the Seismological Society of America* 92: 1025–1031.

Pitarka A, Somerville P, Fukushima Y, Uetake T, and Irikura K (2000) Simulation of near-fault strong ground motion using hybrid Green's functions. *Bulletin of the Seismological Society of America* 90: 566–586.

Pitarka AS, Takenaka H, and Suetsugu D (1994) Modeling strong motion in the Ashigara Valley for the 1990 Odawara, Japan, earthquake. *Bulletin of the Seismological Society of America* 84: 1327–1335.

Power M, Chiou B, Abrahamson N, Bozorgnia Y, Shantz T, and Roblee C (2008) An overview of the NGA project. *Earthquake Spectra* 24: 3–21.

Prejean SG and Ellsworth WL (2001) Observations of earthquake source parameters at 2 km depth in the Long Valley Caldera, eastern California. *Bulletin of the Seismological Society of America* 91: 165–177.

Pulido N and Kubo T (2004) Near-fault strong motion complexity of the 2000 Tottori earthquake (Japan) from a broadband source asperity model. *Tectonophysics* 390(1–4): 177–192.

Purvance MD (2005) *Overturning of Slender Blocks: Numerical Investigation and Application to Precariously Balanced Rocks in Southern California*. PhD Dissertation, University of Nevada, Reno.

Purvance MD and Anderson JG (2003) A comprehensive study of the observed spectral decay in strong-motion accelerations recorded in Guerrero, Mexico. *Bulletin of the Seismological Society of America* 93: 600–611.

Quinay PEB, Ichimura T, and Hori M (2012) Waveform inversion for modeling three-dimensional crust structure with topographic effects. *Bulletin of the Seismological Society of America* 102(3): 1018–1029.

Rautian TC and Khalturin VI (1978) The use of coda for determination of the earthquake source spectrum. *Bulletin of the Seismological Society of America* 68: 923–948.

Richards PG and Menke W (1983) The apparent attenuation of a scattering medium. *Bulletin of the Seismological Society of America* 73: 1005–1021.

Rietbrock A, Strasser F, and Edwards B (2013) A stochastic earthquake ground-motion prediction model for the United Kingdom. *Bulletin of the Seismological Society of America* 103(1): 57–77.

Rodriguez-Marek A, Montalva GA, Cotton F, and Bonilla F (2011) Analysis of single-station standard deviation using the KiK-net data. *Bulletin of the Seismological Society of America* 101(3): 1242–1258.

Rogers AM and Perkins DM (1996) Monte Carlo simulation of peak-acceleration attenuation using a finite-fault uniform-patch model including isochrone and extremal characteristics. *Bulletin of the Seismological Society of America* 86: 79–92.

Roumelioti Z and Beresnev IA (2003) Stochastic finite-fault modeling of ground motions from the 1999 Chi-chi, Taiwan, earthquake: Application to rock and soil sites with implications for nonlinear site response. *Bulletin of the Seismological Society of America* 93(4): 1691–1702.

Roumelioti Z and Kiratzi A (2002) Stochastic simulation of strong-motion records from the 15 April 1979 (M 7.1) Montenegro earthquake. *Bulletin of the Seismological Society of America* 92(3): 1095–1101.

Roumelioti Z, Kiratzi A, and Theodulidis N (2004) Stochastic strong ground-motion simulation of the 7 September 1999 Athens (Greece) earthquake. *Bulletin of the Seismological Society of America* 94(3): 1036–1052.

Rubinstein JL and Beroza GC (2004) Evidence for widespread nonlinear strong ground motion in the MW 6.9 Loma Prieta earthquake. *Bulletin of the Seismological Society of America* 94(5): 1595–1608.

Safarshahi M, Rezapour M, and Hamzehloo H (2013) Stochastic finite-fault modeling of ground motion for the 2010 Rigan earthquake, southeastern Iran. *Bulletin of the Seismological Society of America* 103(1): 223–235.

Sánchez-Sesma FJ and Campillo M (1991) Diffraction of P, SV, and Rayleigh waves by topographic features: A boundary integral formulation. *Bulletin of the Seismological Society of America* 81: 2234–2253.

Sanchez-Sesma FJ, Rodriguez M, Iturraran-Viveros U, et al. (2011) A theory for microtremor H/V spectral ratio: Application for a layered medium. *Geophysical Journal International* 186: 221–225. http://dx.doi.org/10.1111/j.1365-246X.2011.05064.x.

Sandikkaya MA, Akkar S, and Bard P-Y (2013) A nonlinear site-amplification model for the next Pan-European ground-motion prediction equations. *Bulletin of the Seismological Society of America* 103: 19–32.

Sato T, Graves RW, and Somerville PG (1999) Three-dimensional finite-difference simulations of long-period strong motions in the Tokyo metropolitan area during the 1990 Odawara earthquake (MJ 5.1) and the great 1923 Kanto earthquake (MS 8.2) in Japan. *Bulletin of the Seismological Society of America* 89(3): 579–607.

Sato T, Graves RW, Somerville PG, and Kataoka S (1998) Estimates of regional and local strong motions during the great 1923 Kanto, Japan, earthquake (Ms 8.2). Part 2: Forward simulation of seismograms using variable-slip rupture models and estimation of near-fault long-period ground motions. *Bulletin of the Seismological Society of America* 88: 206–227.

Sato T, Helmberger DV, Somerville PG, Graves RW, and Saikia CK (1998) Estimates of regional and local strong motions during the great 1923 Kanto, Japan, earthquake (Ms 8.2). Part 1: Source estimation of a calibration event and modeling of wave propagation paths. *Bulletin of the Seismological Society of America* 88: 183–205.

Sato T, Kawase H, Matsui M, and Kataoka S (1991) Array measurement of high frequency microtremors for underground structure estimation. In: *Proceedings of the 4th International Conference on Seismic Zonation*, Odawana, Japan, vol. 2, pp. 409–416.

Satoh T (2002) Empirical frequency-dependent radiation pattern of the 1998 Miyagiken-Nanbu earthquake in Japan. *Bulletin of the Seismological Society of America* 92(3): 1032–1039.

Satoh T, Fushimi M, and Tatsumi Y (2001) Inversion of strain-dependent nonlinear characteristics of soils using weak and strong motions observed by borehole sites in Japan. *Bulletin of the Seismological Society of America* 91: 365–380.

Satoh T, Kawase H, Sato T, and Pitarka A (2001) Three-dimensional finite-difference waveform modeling of strong motions observed in the Sendai Basin, Japan. *Bulletin of the Seismological Society of America* 91(4): 812–825.

Satoh T, Sato T, and Kawase H (1995) Nonlinear behavior of soil sediments identified by using borehole records observed at the Ashigara valley, Japan. *Bulletin of the Seismological Society of America* 85: 1770–1789.

Savage JC (1972) Relation of corner frequency to fault dimensions. *Journal of Geophysical Research* 77: 3788–3795.

Scherbaum F (2002) Analysis of digital earthquake signals. In: Lee WHK, Kanamori H, and Jennings PC (eds.) *International Handbook of Earthquake and Engineering Seismology*, vol. 81A, pp. 349–367. London: International Association Seismology and Physics of the Earth's Interior, Committee on Education, Academic Press.

Schnabel PB, Lysmer J, and Seed HB (1972) SHAKE: A computer program for earthquake response analysis of horizontally layered sites. *Report EERC 72-12*. Berkeley, CA: University of California.

Schneider JF, Silva WJ, and Stark C (1993) Ground motion model for the 1989 M 6.9 Loma Prieta earthquake including effects of source, path, and site. *Earthquake Spectra* 9: 251–287.

Scholz CH, Aviles CA, and Wesnousky SG (1986) Scaling differences between large intraplate and interplate earthquakes. *Bulletin of the Seismological Society of America* 76: 65–70.

Scognamiglio L, Malagnini L, and Akinci A (2005) Ground-motion scaling in Eastern Sicily, Italy. *Bulletin of the Seismological Society of America* 95: 568–578.

Seale SH and Archuleta RJ (1989) Site amplification and attenuation of strong ground motion. *Bulletin of the Seismological Society of America* 79: 1673–1696.

Sekiguchi H and Iwata T (2001) The source process of the 1999 Chi-Chi, Taiwan, earthquake in semi-long period (2-20s). *Annual Report on Active Fault and Paleoearthquake Researches* 1: 315–324.

Semblat JF, Kham M, and Bard P-Y (2008) Seismic-wave propagation in alluvial basins and influence of site-city interaction. *Bulletin of the Seismological Society of America* 98(6): 2665–2678.

Shakal AF, Huang MJ, and Graizer VM (2003) Strong-motion data processing. In: Lee WHK, Kanamori H, Jennings PC, and Kisslinger C (eds.) *International Handbook of Earthquake and Engineering Seismology*, vol. 81B, pp. 967–981. Amsterdam: International Association of Seismology and Physics of the Earth's Interior, Academic Press.

Shapiro NM, Campillo M, Singh SK, and Pacheco J (1998) Seismic channel waves in the accretionary prism of the Middle America Trench. *Geophysical Research Letters* 25: 101–104.

Shapiro NM, Olsen KB, and Singh SK (2000) Wave-guide effects in subduction zones: Evidence from three-dimensional modeling. *Geophysical Research Letters* 27: 433–436.

Shaw BE (2013) Earthquake surface slip-length data is fit by constant stress drop and is useful for seismic hazard analysis. *Bulletin of the Seismological Society of America* 103: 876–893.

Shi B, Anooshehpoor A, Zeng Y, and Brune JN (1996) Rocking and overturning of precariously balanced rocks by earthquakes. *Bulletin of the Seismological Society of America* 86: 1364–1371.

Shin TC, Tsai YB, Yeh YT, Liu CC, and Wu YM (2003) Strong-motion instrumentation programs in Taiwan. In: Lee WHK, Kanamori H, and Jennings PC (eds.) *International Handbook of Earthquake and Engineering Seismology*, vol. 81B, pp. 1057–1062. London: International Association Seismology and Physics of the Earth's Interior, Committee on Education, Academic Press.

Sibson RH (1984) Roughness at the base of the seismogenic zone: Contributing factors. *Journal of Geophysical Research* 89: 5791–5799.

Silva W, Darragh R, Stark C, et al. (1990) A methodology to estimate design response spectra in the near-source region of large earthquakes using the band-limited-white-noise ground motion model. In: *Proceedings of the Fourth U.S. National Conference on Earthquake Engineering*, vol. 1, pp. 487–494.

Silva W, Wong I, Siegel T, Gregor N, Darragh R, and Lee R (2003) Ground motion and liquefaction simulation of the 1886 Charleston, South Carolina, earthquake. *Bulletin of the Seismological Society of America* 93(6): 2717–2736.

Singh SK, Ordaz M, and Pacheco JF (2003) Advances in seismology with impact on earthquake engineering. In: Lee WHK, Kanamori H, and Jennings PC (eds.) *International Handbook of Earthquake and Engineering Seismology*, vol. 81B, pp. 1081–1095. London: International Association Seismology and Physics of the Earth's Interior, Committee on Education, Academic Press.

Smerzini C and Villani M (2012) Broadband numerical simulations in complex near-field geological configurations: The case of the 2009 Mw 6.3 L'Aquila earthquake. *Bulletin of the Seismological Society of America* 102(6): 2436–2451.

Sokolov VY (1997) Empirical models for estimating Fourier-amplitude spectra of ground acceleration in the Northern Caucasus (Racha seismogenic zone). *Bulletin of the Seismological Society of America* 87(6): 1401–1412.

Somerville P, Irikura K, Graves R, et al. (1999) Characterizing crustal earthquake slip models for the prediction of strong ground motion. *Seismological Research Letters* 70: 59–80.

Somerville P and Moriwaki Y (2003) Seismic hazards and risk assessment in engineering practice. In: Lee WHK, Kanamori H, and Jennings PC (eds.) *International Handbook of Earthquake and Engineering Seismology*, vol. 81B, pp. 1065–1080. London: International Association Seismology and Physics of the Earth's Interior, Committee on Education, Academic Press.

Somerville P, Sen M, and Cohee B (1991) Simulation of strong ground motions recorded during the 1985 Michoacan, Mexico and Valparaiso, Chile earthquakes. *Bulletin of the Seismological Society of America* 81(1): 1–27.

Somerville PG, Smith NF, Graves RW, and Abrahamson NA (1997) Modification of empirical strong ground motion attenuation relations to include the amplitude and duration effects of rupture directivity. *Seismological Research Letters* 68(1): 199–222.

Spudich P and Frazer LN (1984) Use of ray theory to calculate high-frequency radiation from earthquake sources having spatially variable rupture velocity and stress drop. *Bulletin of the Seismological Society of America* 74: 2061–2082.

Spudich P and Frazer LN (1987) Erratum. *Bulletin of the Seismological Society of America* 77: 2245.

Spudich PA, Hellwig M, and Lee WHK (1996) Directional topographic response at Tarzana observed in aftershocks of the 1994 Northridge, California, earthquake: Implications for mainshock motions. *Bulletin of the Seismological Society of America* 86(1B): S193–S208.

Spudich P and Iida M (1993) The seismic coda, site effects, and scattering in alluvial basins studied using aftershocks of the 1986 North Palm Springs, California, earthquake as source arrays. *Bulletin of the Seismological Society of America* 83: 1721–1743.

Steidl JH, Tumarkin AG, and Archuleta RJ (1996) What is a reference site? *Bulletin of the Seismological Society of America* 86: 1733–1748.

Stidham C, Antolik M, Dreger D, Larsen S, and Romanowicz B (1999) Three-dimensional structure influences on the strong-motion wavefield of the 1989 Loma Prieta earthquake. *Bulletin of the Seismological Society of America* 89: 1184–1202.

Stirling MW, Anooshehpoor A, Brune JN, Biasi GP, and Wesnousky SG (2002) Assessment of the site conditions of precariously balanced rocks in the Mojave Desert, southern California. *Bulletin of the Seismological Society of America* 92(6): 2139–2144.

Stork AL and Ito H (2004) Source parameter scaling for small earthquakes observed at the Western Nagano 800-m-deep borehole, central Japan. *Bulletin of the Seismological Society of America* 94: 1781–1794.

Su F, Anderson JG, and Zeng Y (1998) Study of weak and strong ground motion including nonlinearity from the Northridge, California, earthquake sequence. *Bulletin of the Seismological Society of America* 88(6): 1411–1425.

Taborda R and Bielak J (2013) Ground-motion simulation and validation of the 2008 Chino Hills, California, earthquake. *Bulletin of the Seismological Society of America* 103(1): 131–156.

Tajima R and Tajima F (2007) Seismic scaling relations and aftershock activity from the sequences of the 2004 mid Niigata and the 2005 west off Fukuoka earthquakes (Mw 6.6) in Japan. *Journal of Geophysical Research* 112: B10302, http://dx.doi.org/10.1029/2007JB004941.

Takahashi T, Sato H, Ohtake M, and Obara K (2005) Scale dependence of apparent stress for earthquakes along the subducting Pacific plate in northeastern Honshu, Japan. *Bulletin of the Seismological Society of America* 95: 1334–1345.

Takahasi R (1956) The SMAC. Strong motion accelerograph and other latest instruments for measuring earthquakes and building vibrations. In: *Proceedings of the First World Conference in Earthquake Engineering*, Berkeley, CA, vol. 3, pp. 1–11.

Takemura M, Motosaka M, and Yamanaka H (1995) Strong motion seismology in Japan. *Journal of Physics of the Earth* 43: 211–257.

Tape C, Liu Q, Maggi A, and Tromp J (2011) Adjoint tomography of the southern California crust. *Science* 325: 988–992. http://dx.doi.org/10.1126/science.117298.

Thompson EM, Baise LG, Kayen RE, and Guzina BB (2009) Impediments to predicting site response: Seismic property estimation and modeling simplifications. *Bulletin of the Seismological Society of America* 99: 2927–2949. http://dx.doi.org/10.1785/0120080224.

Trifunac MD and Brady AG (1975) A study of the duration of strong earthquake ground motion. *Bulletin of the Seismological Society of America* 65: 581–626.

Trifunac MD and Hudson DE (1971) Analysis of the Pacoima Dam accelerogram, San Fernando, California, earthquake of 1971. *Bulletin of the Seismological Society of America* 61: 1393–1411.

Trifunac MD and Todorovska MI (2001) Evolution of accelerographs, data processing, strong motion arrays and amplitude and spatial resolution in recording strong earthquake motion. *Soil Dynamics and Earthquake Engineering* 21: 537–555.

Tsuda K, Archuleta RJ, and Koketsu K (2006) Quantifying the spatial distribution of site response by use of the Yokohama high-density strong-motion network. *Bulletin of the Seismological Society of America* 96: 926–942.

Tsuda K, Steidl J, Archuleta R, and Assimaki D (2006) Site-response estimation for the 2003 Miyagi-Oki earthquake sequence considering nonlinear site response. *Bulletin of the Seismological Society of America* 96(4A): 1474–1482. http://dx.doi.org/10.1785/0120050160.

Tsujiura M (1978) Spectral analysis of the coda waves from local earthquakes. *Bulletin of the Earthquake Research Institute, Tokyo University* 53: 1–48.

Tucker BE and King JL (1984) Dependence of sediment-filled valley response on input amplitude and valley properties. *Bulletin of the Seismological Society of America* 74: 153–165.

Tucker BE, King JL, Hatzfeld D, and Nersesov IL (1984) Observations of hard rock site effects. *Bulletin of the Seismological Society of America* 74: 121–136.

Tumarkin AG and Archuleta RJ (1994) Empirical ground motion prediction. *Annali di Geofisica* 37(6): 1691–1720.

Tumarkin AG, Archuleta RJ, and Madariaga R (1994) Scaling relations for composite earthquake models. *Bulletin of the Seismological Society of America* 84(4): 1279–1283.

Udwadia FE and Trifunac MD (1973) Damped Fourier spectrum and response spectra. *Bulletin of the Seismological Society of America* 63: 1775–1783.

Ugurhan B, Askan A, Akinci A, and Malagnini L (2012) Strong-ground-motion simulation of the 6 April 2009 L'Aquila, Italy, earthquake. *Bulletin of the Seismological Society of America* 102(4): 1429–1445.

Venkataraman A and Kanamori H (2004) Observational constraints on the fracture energy of subduction zone earthquakes. *Journal of Geophysical Research* 109: B05302. http://dx.doi.org/10.1029/2003JB002549.

Vernon FL, Fletcher J, Carroll L, Chave A, and Sembera E (1991) Coherence of seismic body waves from local events as measured by a small-aperture array. *Journal of Geophysical Research* 96: 11981–11996.

Vidale J (1989) Influence of focal mechanism on peak accelerations on strong motions of the Whittier Narrows, California, earthquake and an aftershock. *Journal of Geophysical Research* 94: 9607–9613.

Vidale JE, Helmberger DH, and Clayton RW (1985) Finite-difference seismograms for SH waves. *Bulletin of the Seismological Society of America* 75: 1765–1782.

Wald DJ and Graves RW (1998) The seismic response of the Los Angeles basin, California. *Bulletin of the Seismological Society of America* 88: 337–356.

Webb S (2002) Seismic noise on land and on the sea floor. In: Lee WHK, Kanamori H, and Jennings PC (eds.) *International Handbook of Earthquake and Engineering Seismology*, vol. 81A, pp. 305–318. London: International Association Seismology and Physics of the Earth's Interior, Committee on Education, Academic Press.

Wells DL and Coppersmith KJ (1994) New empirical relationships among magnitude, rupture length, rupture width, rupture area, and surface displacement. *Bulletin of the Seismological Society of America* 84: 974–1002.

Wennerberg LG (1990) Stochastic summation of empirical Green's functions. *Bulletin of the Seismological Society of America* 80(6): 1418–1432.

Wesnousky SG (1986) Earthquakes, quaternary faults, and seismic hazard in California. *Journal of Geophysical Research* 91: 12587–12631.

Wielandt E (2002) Seismometry. In: Lee WHK, Kanamori H, and Jennings PC (eds.) *International Handbook of Earthquake and Engineering Seismology*, vol. 81A, pp. 283–304. London: International Association Seismology and Physics of the Earth's Interior, Committee on Education, Academic Press.

Wirgin A and Bard P-Y (1996) Effects of buildings on the duration and amplitude of ground motion in Mexico City. *Bulletin of the Seismological Society of America* 86(3): 914–920.

Wu CJ, Takeo M, and Ide S (2001) Source process of the Chi-Chi earthquake: A joint inversion of strong motion data and global positioning system data with a multifault model. *Bulletin of the Seismological Society of America* 91(5): 1128–1143.

Wyss M (1970) Stress estimates for South American shallow and deep earthquakes. *Journal of Geophysical Research* 75(8): 1529–1544. http://dx.doi.org/10.1029/JB075i008p01529.

Xu H, Day SM, and Minster J-BH (1999) Two-dimensional linear and nonlinear wave propagation in a half-space. *Bulletin of the Seismological Society of America* 89(4): 903–917.

Yamada M, Mori J, and Heaton T (2009) The slapdown phase in high-acceleration records of large earthquakes. *Seismological Research Letters* 80: 559–564.

Yang J and Sato T (2000) Interpretation of seismic vertical amplification observed at an array site. *Bulletin of the Seismological Society of America* 90: 275–285.

Yang J, Sato T, Savidis S, and Li XS (2002) Horizontal and vertical components of earthquake ground motions at liquefiable sites. *Soil Dynamics and Earthquake Engineering* 22: 229–240.

Ye L, Lay T, and Kanamori H (2013) Ground shaking and seismic source spectra for large earthquakes around the megathrust fault offshore of northeastern Honshu, Japan. *Bulletin of the Seismological Society of America* 103: 1221–1241. http://dx.doi.org/10.1785/0120120115.

Yomogida K, Aki K, and Benites R (1997) Coda Q in two-layer random media. *Geophysical Journal International* 128(2): 425–433.

Yomogida K and Etgen JT (1993) 3-D wave propagation in the Los Angeles basin for the Whittier-Narrows earthquake. *Bulletin of the Seismological Society of America* 83: 1325–1344.

Yoshimura C, Bielak J, Hisada Y, and Fernández A (2003) Domain reduction method for three-dimensional earthquake modeling in localized regions, Part II: Verification and applications. *Bulletin of the Seismological Society of America* 93: 825–841.

Yu G, Anderson JG, and Siddharthan R (1993) On the characteristics of nonlinear soil response. *Bulletin of the Seismological Society of America* 83: 218–244.

Zeng Y (1991) Compact solutions of multiple scattered wave energy in time domain. *Bulletin of the Seismological Society of America* 81: 1022–1029.

Zeng Y (1993) Theory of scattered P- and S-wave energy in a random isotropic scattering medium. *Bulletin of the Seismological Society of America* 83: 1264–1276.

Zeng Y (2006) Broadband strong ground motion simulation using a composite source model. In: *Proceedings of the 8th U.S. National Conference on Earthquake Engineering*, 18–22 April 2006, San Francisco, CA, Paper No. 1445.

Zeng Y, Anderson JG, and Su F (1995) Subevent rake and random scattering effects in realistic strong ground motion simulation. *Geophysical Research Letters* 22(1): 17–20.

Zeng Y, Anderson JG, and Yu G (1994) A composite source model for computing realistic synthetic strong ground motions. *Geophysical Research Letters* 21(8): 725–728.

Zeng YH and Chen CH (2001) Fault rupture process of the 20 September 1999 Chi-Chi, Taiwan, earthquake. *Bulletin of the Seismological Society of America* 91(5): 1088–1098.

Zeng Y, Su F, and Aki K (1991) Scattered wave energy propagation in random isotropic scattering medium – Part I: Theory. *Journal of Geophysical Research* 96: 607–619.

Zhao JX and Xu H (2013) A comparison of VS30 and site period as site-effect parameters in response spectral ground-motion prediction equations. *Bulletin of the Seismological Society of America* 103(1): 1–18.

This page intentionally left blank

4.21 Paleoseismology

L Grant Ludwig, University of California, Irvine, CA, USA

© 2015 Elsevier B.V. All rights reserved.

4.21.1	**Introduction**	559
4.21.1.1	Paleoseismology and Its Significance	559
4.21.1.2	Development and Applications of Paleoseismology	560
4.21.1.3	Purpose and Scope	560
4.21.2	**Methods**	560
4.21.2.1	Fault Identification	560
4.21.2.1.1	Recognition of tectonic deformation	560
4.21.2.1.2	Application of digital topographic and imaging data	561
4.21.2.1.3	Uncertainty in fault recognition	561
4.21.2.2	Application of Geochronology	561
4.21.2.2.1	Radiocarbon dating	562
4.21.2.2.2	Statistical analysis of geochronology data	563
4.21.2.3	Identifying and Cataloging Paleoearthquakes	563
4.21.2.3.1	The southern San Andreas fault paleoearthquake record	564
4.21.2.3.2	Surface rupture data	564
4.21.2.3.3	Regional coseismic deformation	567
4.21.2.3.4	Ground shaking and secondary effects	568
4.21.2.3.5	Measurement of slip	569
4.21.2.4	Magnitude of Paleoearthquakes	569
4.21.2.5	Slip Rate	570
4.21.3	**Models and Interpretation of Data**	571
4.21.3.1	Segmentation Models	572
4.21.3.2	Characteristic Earthquake Model	573
4.21.3.3	Time-Predictable and Recurrence Models	574
4.21.3.4	Interpretation of Paleoseismic Data	574
4.21.3.4.1	Uncertainty in paleoseismic data and implications for the reliability of seismic hazard assessment	575
Acknowledgments		575
References		575

4.21.1 Introduction

4.21.1.1 Paleoseismology and Its Significance

Earthquakes are among the most powerful and destructive events on earth. Devastating earthquakes offshore Indonesia (26 December 2004), in Haiti (12 January 2010), and offshore Tōhoku, Japan (11 March 2011), collectively killed half a million people and had global impact (Calais et al., 2010; Goldfinger et al., 2013; Tajima et al., 2013). Understanding earthquakes, and the behavior of faults that produce them, is a socially compelling grand challenge of science. Yet the earthquake problem is difficult to address because of its scale and complexity. Large earthquakes affect vast areas and recur on timescales that make it impossible to study them under controlled laboratory conditions, yet population growth in seismically active areas has placed millions at risk. There is a critical need for observational data on the characteristics of large earthquakes, including their surface effects and their spatial and temporal patterns of occurrence. Such data can provide constraints on models of fault behavior, fault interaction, and tectonic processes and insights into the likely impacts of future earthquakes.

The time intervals between large earthquakes range from decades to many thousands of years depending on the slip rate and tectonic environment of the fault (Sibson, 2002). Thus, the modern scientific observation of earthquakes and the instrumental data set cover only a fraction of relevant observation time for an individual fault or region. Paleoseismology is a subdiscipline of seismology that provides observational data on large earthquakes over millennial timescales through the application of geologic methods (Grant, 2002). The observations are typically described as *paleoearthquakes* or *paleoseismic events* or *paleoseismic event evidence* to distinguish from earthquakes cataloged with modern instrumental methods. Paleoseismic data sets may provide the only observations of extreme events in a region or of the largest earthquakes on a specific fault (Goldfinger et al., 2013).

Paleoseismic data include observations of prehistoric paleoearthquakes and preinstrumental earthquakes. Both types of observations supplement the relatively short historic and instrumental record of earthquakes (McCalpin and Nelson, 1996; Sieh, 1981; Wallace, 1981). Because the primary methods of paleoseismic data collection and analysis are drawn from geology, paleoseismology is sometimes considered to be a subspecialty of active tectonics or earthquake geology (McCalpin, 2009; Yeats et al., 1997). Some methods employed in paleoseismic research, such as hand excavation and

examination of late Holocene sediments, are similar to archaeological field methods. Archaeoseismology is a subdiscipline of paleoseismology or archaeology that uses archaeological methods to collect and interpret evidence of paleoearthquakes (Nur, 2007).

4.21.1.2 Development and Applications of Paleoseismology

Paleoseismology is a young science that emerged as a distinct discipline in the late 1960s to mid-1980s. The history of the discipline can be traced through publication of summaries, compendia, or review articles such as Wallace (1981), Sieh (1981), Wallace (1986), Crone and Omdahl (1987), Vittori et al. (1991), Pantosti and Yeats (1993), Prentice et al. (1994), Serva and Slemmons (1995), Yeats and Prentice (1996), Pavlides et al. (1999), Grant and Lettis (2002), and Grant (2002, 2007). Yeats et al. (1997) published a textbook on the geology of earthquakes that integrates paleoseismology with global coverage of active tectonics. A comprehensive text and reference book on paleoseismology (McCalpin, ed.) was first published in 1996. The second edition (McCalpin, 2009) remains a standard of reference for paleoseismology researchers and practitioners in North America.

Paleoseismic data can be applied toward understanding the behavior of faults, estimating the probability and impacts of future earthquakes, and for designing structures to mitigate seismic risk. Paleoseismic data are used for seismic source characterization in hazard assessment (Field et al., 2009; Reiter, 1991, 1995) and mitigation (Somerville and Moriwaki, 2002). Probabilistic seismic hazard assessment and calculation of fault rupture probability (e.g., Field et al., 2009; Petersen et al., 2008) require paleoseismic information on fault location and geometry, slip rates, recurrence intervals, dates of previous ruptures, and maximum earthquakes (McCalpin, 1996; Yeats et al., 1997). Deterministic hazard assessments use paleoseismic data to develop earthquake scenarios including rupture location, magnitude, and surface displacement. Paleoseismic data can be used to develop earthquake scenarios (e.g., Jones et al., 2008) that are useful for emergency response planning or rupture scenarios that are useful for calculating rupture probabilities (Biasi and Weldon, 2009). Maximum magnitude earthquakes for a fault are estimated from segmentation models, slip per event, estimated lengths of previous ruptures, and their signature in the geologic record (Biasi and Weldon, 2006; Goldfinger et al., 2013). Surface rupture hazard can be mitigated by identifying fault zones that are most likely to rupture and restricting land uses within such zones. Paleoseismic investigations provide data on the recency and frequency of fault rupture and the type and width of surface rupture. This information can be used to classify fault zones according to rupture potential and expected displacement in future earthquakes (e.g., California Geological Survey, 2002). Critical structures can be sited to avoid rupture hazard zones (McCalpin, 1996; Mualchin, 2005) or designed to accommodate expected displacement. The establishment of digital databases (Grant et al., 2005) and standard formats for paleoseismic data (Grant and Gould, 2004) allows better integration with more quantitative fields of seismology, seismic hazard assessment, and earthquake engineering.

4.21.1.3 Purpose and Scope

The purpose of this chapter is to summarize methods, contributions, and issues in paleoseismology. This chapter is an expanded and updated version of earlier tutorial and review chapters (Grant, 2002, 2007). Topics include methods for collecting and analyzing paleoseismic data and interpretation of results. Examples and references are chosen to highlight major concepts and findings rather than attempting to provide a comprehensive or geographically inclusive summary. Most examples are from western North America, and heavy emphasis is placed on the studies of the San Andreas fault system. The chapter concludes with a discussion of issues in interpreting paleoseismic data and applying them to understanding fault systems or constraining seismic hazard.

4.21.2 Methods

With few exceptions, earthquakes are generated by the movement of faults (Bolt, 1999). Therefore, paleoseismic research is directly or indirectly a study of faults and their surface expression (Grant, 2002). Moderate to large earthquakes may generate surface rupture or induce permanent changes in the landscape and local environment. Under suitable conditions, evidence of earthquakes is preserved in the geologic record. *Paleoearthquakes* are recognized by detailed observations and analyses of geologic or environmental conditions within fault zones or in tectonically or seismically active regions.

Paleoseismic investigation methods employ techniques from several geologic subdisciplines, including stratigraphic analysis, Quaternary geology, soil science, geomorphology, engineering geology, geochronology, and structural geology. Major objectives of paleoseismic studies include the identification of recently active or seismogenic faults, measurement of displacement (coseismic and long-term average), and establishment of catalogs of paleoearthquakes. This section is organized by common goals of paleoseismic studies, rather than by disciplinary methods, because most investigations employ a variety of techniques.

4.21.2.1 Fault Identification

Because most earthquakes are generated by faults, the first step in most paleoseismic investigations is the identification of a fault for study. An active fault is a fault that may have displacement within a future period of concern to humans (Wallace, 1981). The identification of active faults requires recognizing previous displacement and constraining the age of displacement. Therefore, an active fault is usually identified by associating it with tectonically deformed Quaternary-age materials or surfaces (e.g., Ziony and Yerkes, 1985) such as the offset stream channels and linear landforms shown in **Figure 1**.

4.21.2.1.1 Recognition of tectonic deformation

Most paleoseismic investigation methods focus on faults that cause recognizable tectonic deformation or environmental changes (e.g., submergence, erosion, anomalous deposition, or mortality) at or near the Earth's surface. Faults may be classified by average slip rate (Sibson, 2002), type of displacement

Figure 1 Photograph of the active trace of the San Andreas fault in the Carrizo Plain, CA, shows classic geomorphology of strike-slip faults, including a (dry) sag pond depression and right-laterally offset stream channels and gullies. The most recent surface rupture occurred during the Fort Tejon earthquake of 9 January 1857. An alluvial fan on the former Phelan Ranch, visible in the lower third of the photo, was offset ~6.7 m in 1857 and subsequently buried with fresh sediment (Grant and Sieh, 1993). Reproduced from Wallace RE (ed.) (1990) The San Andreas Fault System, California. *U.S. Geological Survey Professional Paper 1515*.

(strike-slip, dip-slip normal, dip-slip reverse, thrust, or oblique-slip), and depth (surface or blind). The depth, rate, and type of fault slip determine the amount and nature of fault-induced deformation at the surface. The faster the slip rate, the more active the fault and the more likely the fault can be recognized and characterized by paleoseismologists (Slemmons and dePolo, 1986). Faults that reach the surface are easier to recognize than blind faults, but they may rupture to the surface infrequently if they have low slip rates (Sibson, 2002).

4.21.2.1.2 Application of digital topographic and imaging data

Recent developments in airborne and terrestrial digital imaging methods (see **Figure 2**) have revolutionized research methods in paleoseismology, earthquake geology, and active tectonics. The key contribution of these methods is to enable rapid and remote identification of tectonically altered surficial landforms such as subtle fault scarps and offset drainages at scales varying from a few meters to regional. For example, Bevis et al. (2005) used LiDAR (Light Detection And Ranging) instrumentation to produce high-resolution (~1 m) digital elevation models (DEMs) of topography along the San Andreas fault system in southern California, United States. The data set was named 'B4' because it was intended to provide baseline data before an earthquake, to allow precise measurement of slip afterward. However, the B4 data have provided important insight about past ruptures and allowed the identification of previously overlooked fault strands and subtly offset features (Arrowsmith and Zielke, 2009; Zielke et al., 2010, 2012) as shown in **Figure 3**. The applications of these methods to paleoseismology and fault identification are described by Haddad et al. (2012), and resources for active tectonic and paleoseismology researchers are freely available through http://www.opentopography.org/ 'a portal to high-resolution topography data and tools' funded by the US National Science Foundation.

4.21.2.1.3 Uncertainty in fault recognition

Paleoseismic data provide an incomplete record of earthquakes and fault rupture. Not all earthquakes leave recognizable evidence in the geologic record. Tectonic deformation due to fault slip may be difficult to recognize if rupture does not reach the surface, especially in intraplate regions (e.g., Lettis et al., 1997). Earthquakes on blind faults may induce subtle surface deformations over broad areas (Stein and Yeats, 1989), subsurface sediment deformation (Leon et al., 2009), or distributed evidence of shaking in disrupted sediments (Reiter, 1995). The existence of an active blind fault can be detected by geomorphic analysis (e.g., Bullard and Lettis, 1993; Keller and Pinter, 1996), by mapping Quaternary sediments and identifying areas of uplift (e.g., Grant et al., 1999, 2002), or by recognizing the effects of shaking that are not associated with known surficial faults (Brune, 2002; Reiter, 1995).

The effects of tectonic deformation can be enhanced or obscured by anthropogenic or natural processes such as weathering and erosion. If tectonic changes cannot be readily distinguished from changes induced by other mechanisms, then the results of paleoseismic studies may be biased toward higher or lower estimates of tectonic displacement. Such problems have been noted in a variety of tectonic and human environments (e.g., Nur, 2007; Obermeier, 2009; Ricci Lucchi, 1995; Zachariasen et al., 1999).

4.21.2.2 Application of Geochronology

Geochronology is the dating of earth materials, surfaces, and processes. Geochronology is essential for paleoseismology because it constrains dates of paleoearthquakes and average rates of fault displacement. The most useful geochronological methods for paleoseismic investigations yield high-resolution ages for common, late Quaternary materials such as soils or buried flora and fauna (Lettis and Kelson, 2000). **Table 1** (from Noller et al., 2000a) summarizes geochronological methods for dating Quaternary materials in fault zones. A detailed discussion of this complex topic is beyond the scope of this chapter. A comprehensive compilation and summary of Quaternary geochronological methods and applications is provided by Noller et al. (2000b). The most commonly used methods are also described by Yeats et al. (1997).

All methods listed in **Table 1** have limitations and uncertainties. Uncertainty in geochronological methods is a major

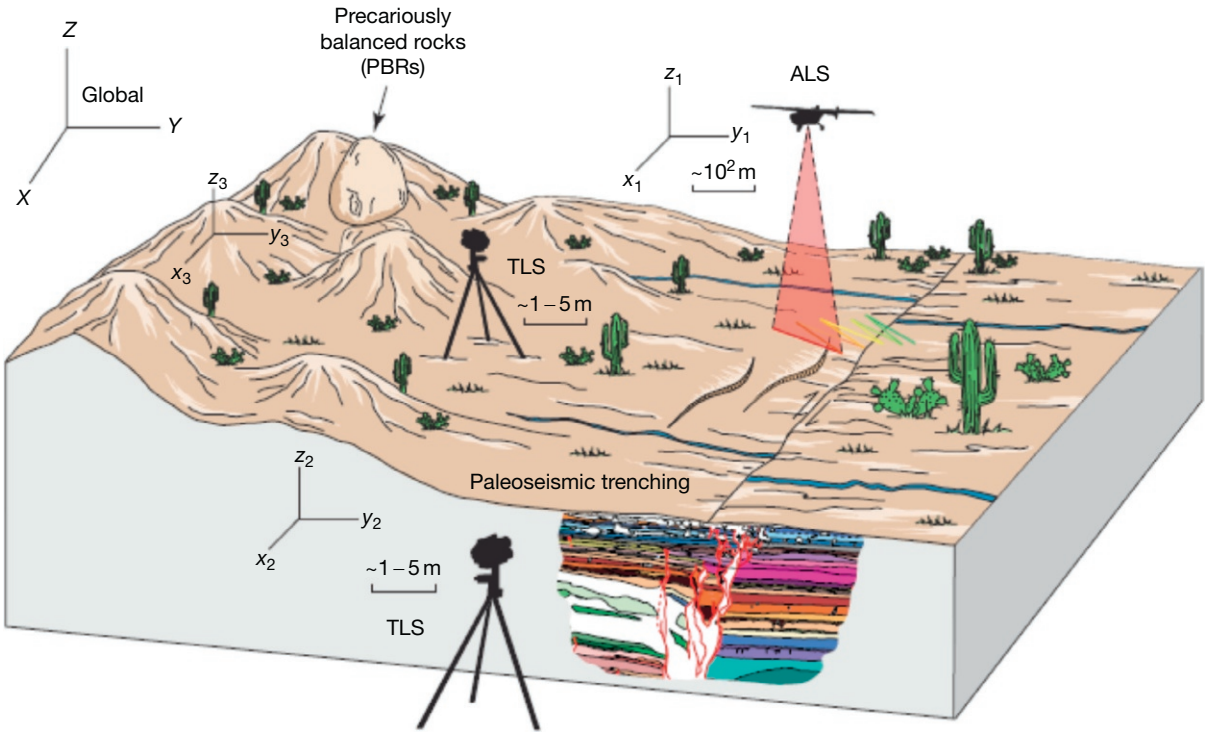


Figure 2 Diagram illustrating the applications of airborne laser scanning (ALS) and terrestrial laser scanning (TLS) for the detection and analysis of earthquake-related deformation. ALS and TLS methods are commonly referred to as LiDAR (Light Detection And Ranging) imaging. ALS-derived digital topographic data illuminate meter-scale fault-related structures and geomorphic elements of fault topography. TLS-derived digital topographic data complement ALS data sets by illuminating fault- and landscape-related components at centimeter to millimeter scales. Reproduced from Haddad DE, Akciz SO, Arrowsmith JR, et al. (2012) Applications of airborne and terrestrial laser scanning to paleoseismology. *Geosphere* 8(4): 771–786, http://dx.doi.org/10.1130/GES00701.1. Themed issue on Seeing the True Shape of Earth's Surface: Applications of Airborne and Terrestrial Lidar (Frankel volume).

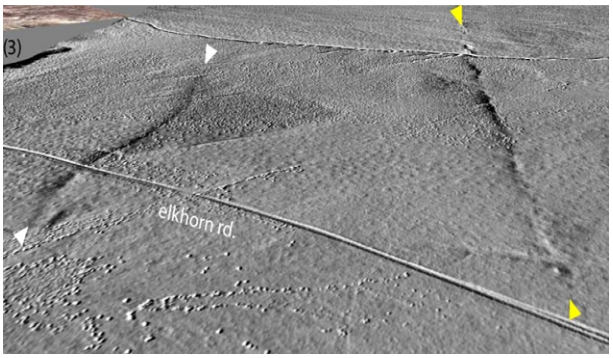


Figure 3 Digital elevation model (DEM) of the San Andreas fault zone in the Carrizo Plain, CA, derived from the B4 data set of Bevis et al. (2005). View to southeast. The main trace of the fault, marked with yellow triangles, has been mapped in detail and studied for several decades. A surface rupture trace, marked with white triangles, was first identified from LiDAR images and subsequently investigated and confirmed to be a fault in 2012 (Grant Ludwig et al., 2013). Image created by Sinan Akciz.

source of uncertainty in paleoseismic data (Lettis and Kelson, 2000). In addition to uncertainty in calibration and accuracy of analysis, errors may be introduced in the selection, collection, and interpretation of field samples. Therefore, most paleoseismic studies employ multiple methods of dating to reduce uncertainty and cross-check results. Methods such as radiocarbon dating that yield accurate, high-precision ages of common or widely distributed materials are preferred. Improvements in geochronological methods and in the statistical treatment of dates have reduced uncertainties in previously published ages (e.g., Akciz et al., 2009; Biasi and Weldon, 1994; Biasi et al., 2002; Sieh et al., 1989).

Recent developments in cosmogenic surface exposure dating methods and analysis of pedogenic carbonate are beginning to be widely applied in paleoseismology to determine the age of offset features for slip rate measurements (e.g., Behr et al., 2010; Fletcher et al., 2010) or to estimate the time since fault scarps have been exposed by recent slip or fragile geologic features have been exposed without toppling (Balco et al., 2011; Bell et al., 1998).

4.21.2.2.1 Radiocarbon dating

Radiocarbon dating is the most widely used method for dating Holocene and latest Pleistocene earthquakes. The half-life of radioactive ^{14}C (5730 years) limits the application of radiocarbon dating to organic matter formed from carbon fixed within the last 50 000–60 000 years (Trumbore, 2000). The amount of ^{14}C in atmospheric CO_2 has varied in the past, particularly in the last few centuries due to anthropogenic emissions. To compensate for this variation, radiocarbon ages are calibrated to correspond to calendar ages (absolute ages). Calibrations

Table 1 Classification of Quaternary geochronological methods

Sidereal	*Isotopic*	*Radiogenic*	*Chemical and biologic*	*Geomorphic*	*Correlation*
Dendrochronology	Radiocarbon	Fission track	Amino acid racemization	Soil-profile development	Stratigraphy
	Uranium series	Thermoluminescence	Rock-varnish cation ratio	Rock-varnish development	Paleomagnetism
Varve chronology	^{210}Pb	Optically stimulated luminescence	Obsidian and tephra hydration	Scarp morphology and landform modification	Tephrochronology
	U–Pb Th–Pb				Paleontology
Historical records		Electron spin resonance	Soil chemistry	Rate of deformation	Tektites
	K–Ar and ^{39}Ar–^{40}Ar			Rate of deposition	Climate correlation
Sclerochronology and growth rings	^{10}Be and other cosmogenic isotopes	Infrared-stimulated luminescence	Lichenometry	Rock and mineral weathering	Astronomical correlation
					Stable isotopes
				Geomorphic position	Archaeology

Adapted from Noller JS, Sowers JM, Colman SM, and Pierce KL (2000) Introduction to quaternary geochronology. In: Noller JS, Sowers JM, and Lettis WR (eds.) *Quaternary Geochronology: Methods and Applications*. AGU Reference Shelf, Series, vol. 4, pp. 1–10. Washington, DC: American Geophysical Union

based on tree rings and glacial varves cover nearly the entire range of ^{14}C dating. Calibration curves are not linear. Plateaus in the calibration curves limit the precision of radiocarbon dating (Trumbore, 2000). This problem is acute for the last few centuries. For example, Yeats and Prentice (1996) noted that the two largest historic ruptures of the San Andreas fault in California, which occurred in 1857 and 1906, are indistinguishable using radiocarbon dating.

4.21.2.2.2 Statistical analysis of geochronology data

In the last decade, paleoseismologists have experimented with the application of Bayesian statistical methods to refine the ages of paleoearthquakes, typically with results of radiocarbon dating (e.g., Biasi and Weldon, 1994; Biasi et al., 2002), leading to improvements in correlation of radiocarbon-dated ruptures between multiple sites on a fault (Weldon et al., 2005). The results of Bayesian analysis is a probability density function (pdf) of ages for a given event based on the pdf's of ^{14}C dates on individual samples below and above the event horizon. Biasi and Weldon (1994) first applied stratigraphic ordering and Bayesian statistics to reduce uncertainty in ages of paleoearthquakes constrained by radiocarbon dating. Biasi and Weldon (2009) extended Bayesian methods to explore the probability that a rupture at different paleoseismic research sites along the same fault could be the same earthquake (**Figure 4**). Ramsey and Lee (2013) provided a popular program, OxCal, that implements Bayesian techniques using stratigraphic ordering and radiocarbon ages to constrain the dates of events. The program is useful for archaeology and paleoseismology. Lienkaemper and Bronk Ramsey (2009) published an OxCal tutorial specifically for paleoseismologists. High-resolution radiocarbon dating methods combined with Bayesian and other statistical methods have led to more quantitative analysis of rupture patterns for the southern San Andreas fault (Akciz et al., 2009, 2010; Scharer et al., 2010, 2011) than was possible with traditional methods.

4.21.2.3 Identifying and Cataloging Paleoearthquakes

Large earthquakes create features that may be preserved in the stratigraphic record and recognized by geologic, geomorphic, environmental, or archaeological analysis. Major effects of earthquakes include surface rupture, ground deformation, liquefaction, and ground failures due to shaking. Chronologies of earthquakes are developed by identifying and dating evidence of multiple paleoearthquakes along a specific fault or by documenting evidence of shaking or tectonic deformation in a seismically active region. Possibly, the longest continuous record of paleoearthquakes spans ~50 000 years, with a mean recurrence time of ~1600 years (Marco et al., 1996).

Nonseismic processes can create features that appear very similar to those generated by earthquakes (Obermeier, 2009; Ricci Lucchi, 1995). This is especially true in coastal areas where vertical tectonic motion is difficult to discriminate from changes in sea level (Zachariasen et al., 1999, 2000) and where offshore turbidite flows can be triggered by earthquakes or large storms (Goldfinger et al., 2012).

The most commonly reported and least ambiguous paleoseismic data are derived from observations of surface ruptures. Well-studied strike-slip faults such as the southern San Andreas fault (Akciz et al., 2010; Grant and Lettis, 2002; Grant Ludwig et al., 2010; Scharer et al., 2010, 2011; Sieh et al., 1989; Weldon et al., 2004, 2005) and the North Anatolian fault (e.g., Hartleb et al., 2003, 2006) have chronologies of a dozen or more paleoearthquakes. The North Anatolian fault ruptured most recently in 1999, culminating a westward propagating sequence of eight $M \geq 7$ earthquakes since AD 1939 (Hartleb et al., 2003). The historic record of earthquakes in Turkey spans approximately 2000 years in the western part of the country (Hartleb et al., 2003; Klinger et al., 2003) near the 1999 rupture. A 2500-year-long record of paleoearthquakes (Hartleb et al., 2006) in north-central Turkey reveals evidence for the five most recent historic earthquakes and at least two

prehistoric paleoearthquakes. The compilation of historic and paleoseismic data suggest that the North Anatolian fault typically ruptures in large, infrequent earthquakes with interevent times that range from 200 to 900 years and that the entire length of the fault ruptured at least twice historically, in sequences similar to the 1939–1999 rupture pattern (Hartleb et al., 2006). Repeated, similar ruptures have occurred along the western segment of the North Anatolian fault (Klinger et al., 2003).

4.21.2.3.1 The southern San Andreas fault paleoearthquake record

The San Andreas fault in California, United States, has perhaps the best documented record of paleoearthquakes of any fault in the world (Grant and Lettis, 2002). Descriptions of large-magnitude historic ruptures in AD 1906 and 1857 are supplemented by the record of paleoearthquakes at many sites along the ~1000 km long fault. The longest and most detailed paleoearthquake records are from the Wrightwood and Pallett Creek sites in southern California (Scharer et al., 2010, 2011; Weldon et al., 2004, 2005). The Wrightwood site has ruptured, on average, approximately once per century since AD 500, although the time interval between earthquakes has varied from ~31 to 165 years (Weldon et al., 2004). Slip per event has varied from 0.7 to 6.6 m, in a complex pattern of strain accumulation and release (Weldon et al., 2004). For the southern San Andreas fault, the paleoearthquake record (see Figure 4) can be fit by models of repeated long ruptures, randomly occurring ruptures of variable length, alternating long and short ruptures (Weldon et al., 2005), or quasiperiodic ruptures (Scharer et al., 2010).

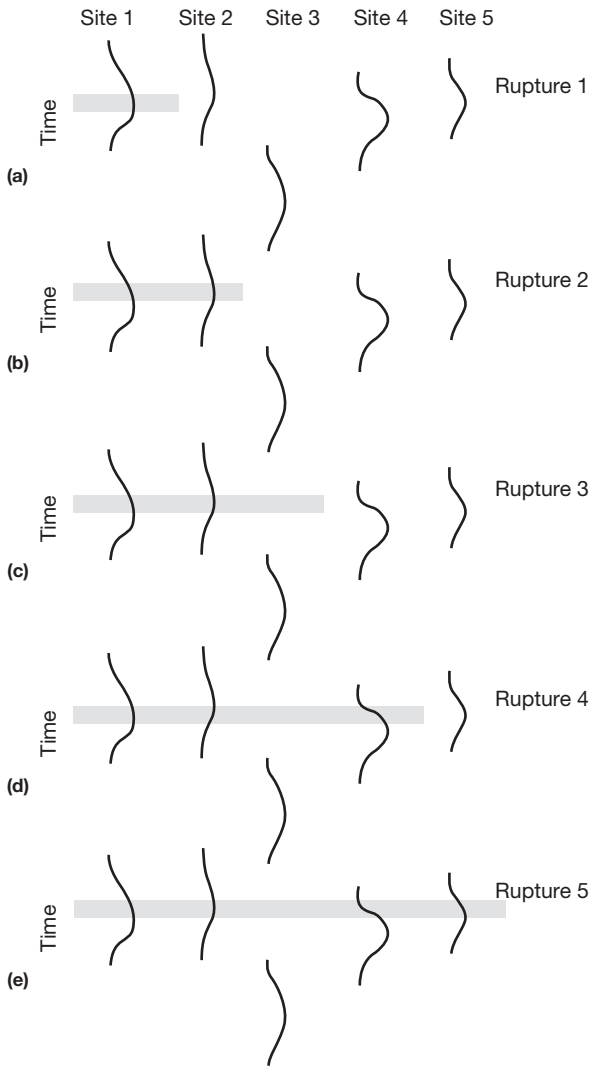


Figure 4 Example of Biasi and Weldon's method of 'stringing pearls' to construct models of fault ruptures (gray horizontal bars) based on paleoevent data from physically separated sites (a) through (e). At each site, paleoearthquake ages are constrained by a pdf of radiocarbon ages (black curves). Reprinted from Biasi GP and Weldon RJ (2009) San Andreas fault rupture scenarios from multiple paleoseismic records: Stringing pearls. *Bulletin of the Seismological Society of America* 99: 477; Copyright: Seismological Society of America.

4.21.2.3.2 Surface rupture data

Evidence of surface rupture can be observed directly in natural exposures or, more commonly, in trench excavations (Figure 5). Paleoseismologists look for trench sites with distinct evidence of multiple ruptures and abundant material for dating at a site with a sedimentation rate rapid enough to bury evidence for each earthquake before the next earthquake occurs. Trenches are generally excavated with a backhoe or similar excavation equipment. Shallow trenches in unconsolidated sediments may be dug by hand. Deep trenches must be stabilized with shoring, benching, or sloping. Stabilization methods are chosen to maximize wall exposure. The maximum practical depth for typical trench investigations is 5 m.

Trench exposures are cleaned by scraping and brushing to highlight stratigraphy and faults. Geologists map the exposures

Figure 5 Photograph of trench wall exposure showing a fault in unconsolidated alluvial fan sediments in the Carrizo Plain. The trench is held open and stabilized by vertical aluminum shores, visible on the right side of the photograph. A string grid (white vertical line on the left side of photo) was installed at 1 m spacing to assist with digital photography and logging. A fault is visible in the middle of the photograph, between the string and shore. Photograph by Sinan Akciz.

with the aid of a string grid or survey control. The maps are called *trench logs*. The goal of logging is to recognize and record the number of surface ruptures and their stratigraphic position relative to datable material so that each paleoearthquake can be analyzed and dated as accurately as possible.

The evidence of prehistoric surface rupture is commonly called a *paleoseismic event*. The stratigraphic unit that was the ground surface at the time of the earthquake is referred to as the *event horizon*. **Figure 6** shows typical stratigraphic and structural relationships that are used for recognizing and dating previous earthquakes (Lettis and Kelson, 2000). **Figure 7** shows an example of a hand-drawn trench log from the San Andreas fault, with several paleoseismic events and event horizons marked with letters. **Figure 8** shows an example of a trench

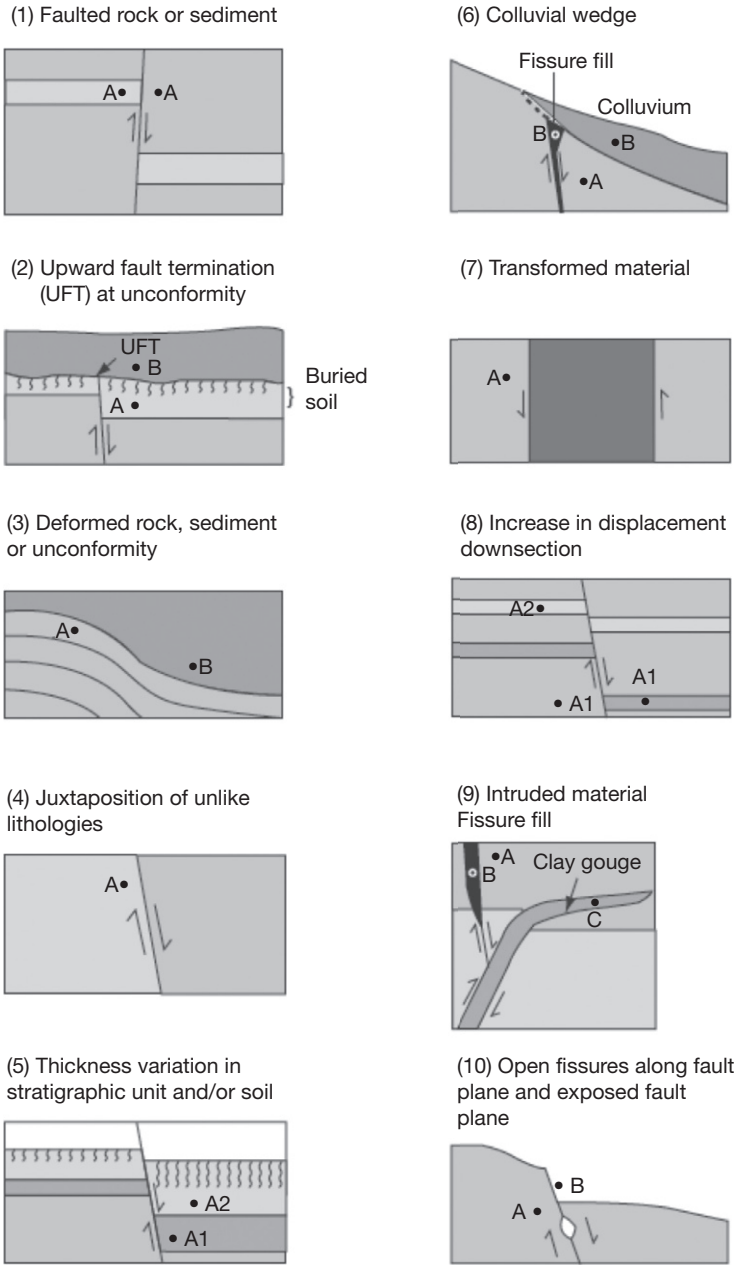


Figure 6 Diagrams illustrating stratigraphic and structural criteria used to identify the occurrence and timing of paleoearthquakes. Stratigraphy, faults, and soils are shown schematically. Dated samples at locations marked 'A' predate the earthquake. Dated samples at locations 'B' postdate the earthquake. Dated samples at 'C' are not helpful for deciphering the chronology of past earthquakes because the sample may predate or postdate the earthquake. Where evidence of multiple earthquakes is present, locations A1 predate the earlier event and locations A2 postdate the earlier event and predate the later event. Reproduced from Lettis WR and Kelson KI (2000) Applying geochronology in paleoseismology. In: Noller JS, Sowers JM, and Lettis WR (eds.) *Quaternary Geochronology: Methods and Applications*. *AGU Reference Shelf, Series*, vol. 4, pp. 479–495. Washington, DC: American Geophysical Union.

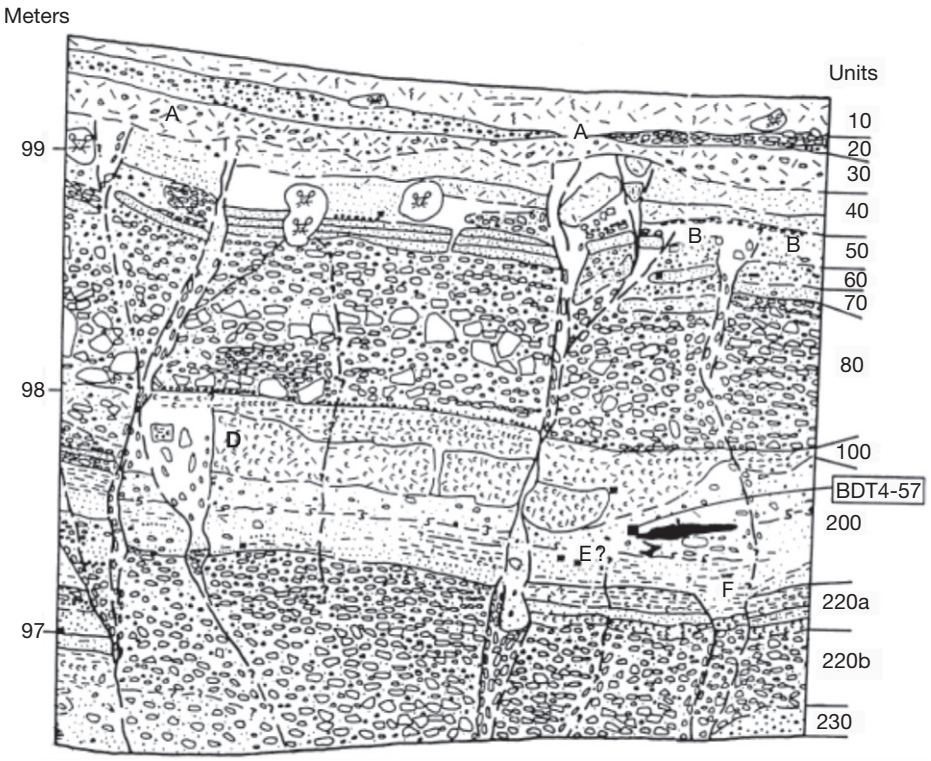


Figure 7 Example of a portion of a trench log from the San Andreas fault at Bidart Ranch in the Carrizo Plain, California. Letters indicate the approximate location of evidence for several paleoearthquakes (A, B, D, E?, F). Query indicates uncertainty. Modified from Grant LB (1993) *Characterization of Large Earthquakes on the San Andreas Fault in the Carrizo Plain: Implications for Fault Mechanics and Seismic Hazard*. Dissertation. California Institute of Technology, Pasadena, CA.

log created with digital methods. To prepare digital logs, the trench walls are digitally photographed (**Figure 9**) or imaged by a terrestrial laser scanner (Haddad et al., 2012). Trench stratigraphy and faults are subsequently logged directly onto photographic or laser-scanned digital images while in the field.

The most useful criteria for identifying ruptures include multiple fault terminations at a single stratigraphic horizon, tilted and folded strata overlain by less deformed strata, and colluvial wedges draped over fault scarps (Yeats and Prentice, 1996). Fault strands may die out and terminate below the surface or be poorly expressed at the event horizon (Bonilla and Lienkaemper, 1990). Therefore, multiple exposures of a fault should be examined whenever possible to identify the event horizon with maximum confidence. Trench logs are supplemented with descriptive notes of observations and objective records such as photographs. Samples of datable material are collected after their locations have been documented on the trench log (**Figure 10**). A good trench log documents observations of stratigraphy and structure as objectively as possible. However, it is impossible to make an entirely objective log because construction of a log is partly an act of interpretation.

In some areas, a fault cannot be examined with excavations because of a high water table, the presence of human structures, or burial by younger sediments. Subsurface investigation methods such as cone penetrometer test (CPT) and large-diameter borings may be applied for these conditions (Dolan et al., 2000; Grant et al., 1997). These methods must be used in combination with other subsurface investigation methods such as seismic imaging, ground-penetrating radar (GPR), and standard borings. There are greater uncertainties in the identification and dating of paleoearthquakes from subsurface investigations than from near-surface exposures because subsurface exposures are either discontinuous, in the case of borings and CPTs, or indirect in the case of seismic or GPR imaging.

The set of surface ruptures revealed by paleoseismic methods is a subset of the actual earthquake history. The paleoseismic record can be spatially and temporally incomplete because the conditions required for preservation of earthquakes are not present at all times on all active faults. The minimum earthquake magnitude associated with surface rupture is about M_L 5 (Bonilla, 1988), so $M_L < 5$ earthquakes are rarely recognized in trench exposures. Surface faulting is commonly associated with $M > 6$ earthquakes (Wells and Coppersmith, 1994). Therefore, most paleoseismic events recognized in trenches represent $M \geq 6$ earthquakes. Not all faults have been identified or studied by paleoseismologists. For example, the devastating 2010 M_w 7 Haiti earthquake occurred on a previously unmapped fault (Calais et al., 2010) rather than the Enriquillo–Plantain Garden fault that was targeted by paleoseismic investigation (Prentice et al., 2010). Early paleoseismic studies focused on examining evidence of larger earthquakes ($M > 7$) because the data were relatively easy to interpret. Therefore, the published paleoseismic record of earthquakes contains less data on the occurrence of smaller ($M < 7$) earthquakes (McCalpin, 1996).

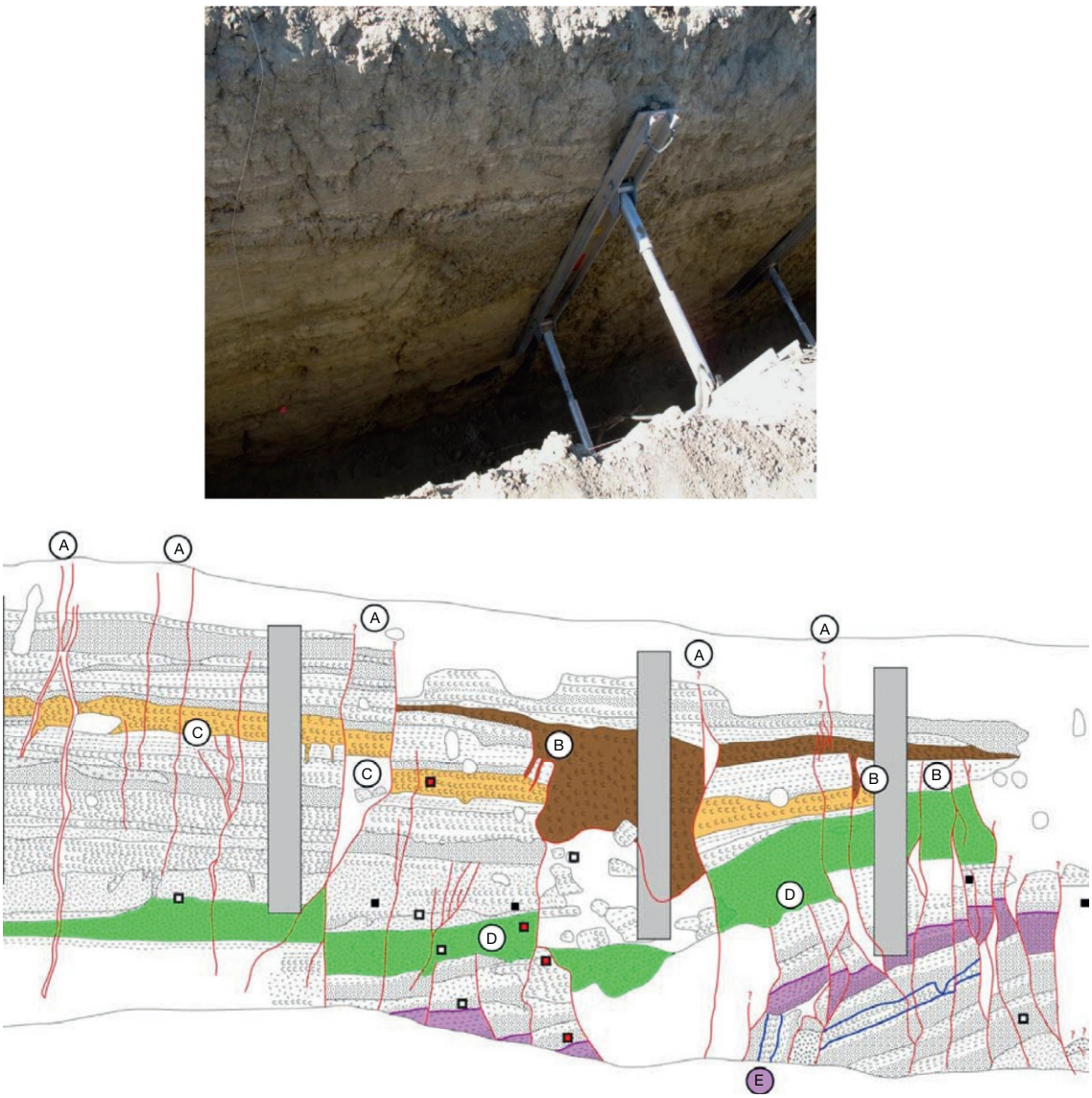


Figure 8 Example photograph showing part of a trench wall (above) and digital log of the same wall (below). The trench is braced with aluminum shores for safety, shown in gray on the digital log. The colors and patterns on the digital log indicate different stratigraphic units. Each unit was identified in the trench and logged on a digital photographic cross section while in the field. A clay-rich unit behind the middle shore is clearly visible on the photo. This unit is colored brown on the digital log. The digital files can be conveniently edited in the office and prepared for publication in summary formats such as this colored version. Photograph by Sinan Akciz.

4.21.2.3.3 Regional coseismic deformation

Earthquakes that generate regional deformation are most likely to be recognized and cataloged if the deformation is vertical. Vertical deformation can induce changes in local rates of deposition and erosion that provide evidence of a paleoearthquake, particularly in fluvial (Schumm et al., 2000) and coastal environments (Carver and McCalpin, 1996). Indicators of coseismic uplift in coastal areas include uplifted or submerged shorelines, platforms or flights of terraces above mean sea level, and unusual growth or die-off patterns in coral microatolls (Zachariasen et al., 1999, 2000). The formation of elevated terraces or submerged shorelines has been reported following historic earthquakes in coastal regions (e.g., Plafker and Rubin, 1978; Matsuda et al., 1978; Pirazzoli, 1991) such as the great Ache–Andaman earthquake of 2004 off the coast of Sumatra (Meltzner et al., 2010). Sequences of similar terraces and shorelines at higher elevations provide evidence of multiple episodes of tectonic uplift. In shallow tropical waters, such as the coastal areas of Sumatra, the growth of corals responds to uplift or subsidence relative to sea level and may preserve a record of vertical coseismic motion (e.g., Natawidjaja et al., 2006; Zachariasen et al., 2000) and aseismic tectonic deformation (Natawidjaja et al., 2004).

Tectonic subsidence induces sedimentation that can record and preserve evidence of an earthquake (Carver and McCalpin, 1996). Buried flora or fauna may provide material for dating

Figure 9 Researchers photograph the walls of a trench in the San Andreas fault zone using a digital camera mounted on a frame that holds the camera level at a fixed distance from the wall for minimal distortion. The photos are 'stitched' together to create a digital log of the trench walls and printed out for manual logging, or mapping, of faults, stratigraphy, and sample locations.

Figure 10 Photo of Dr. Sinan Akciz collecting samples for radiocarbon dating sediments in the San Andreas fault zone on the Bidart Fan in the Carrizo Plain.

the earthquake. For example, in the Cascadia region of northwestern United States, tsunami deposits, drowned forests, and grasslands along the coastline provided material for high-precision dating of subduction zone earthquakes (e.g., Atwater et al., 1995). The date of the most recent earthquake (AD 1700) was resolved from radiocarbon dating of submerged deposits, analysis of tree rings (dendroseismology), and historic accounts of tsunami recorded in Japan (Nelson et al., 1995; Satake et al., 1996).

Measurement and interpretation of vertical deformation in coastal regions must be based on knowledge of sea level at the time of uplift. Sea level was 100–150 m below modern levels during the last glacial maximum 15 000–20 000 years BP (Lajoie, 1986). After rising rapidly in the early Holocene, sea level stabilized in the mid-Holocene. However, sea level has fluctuated slightly (+2 m) during the late Holocene (Lajoie, 1986) and is at different levels in different parts of the world (Pirazzoli, 1991). Fluctuations and regional variations have been sufficiently large to complicate interpretations of coseismic uplift or subsidence in areas with low vertical deformation rates.

4.21.2.3.4 Ground shaking and secondary effects

Surface rupture and regional deformation are considered primary effects of earthquakes because they can be directly associated with tectonic movement of specific faults. Secondary effects of earthquakes such as liquefaction, ground failure, slope failure, tsunami, and seismic seiche are induced by shaking. A history of seismic shaking is an indirect history of regional fault activity and is useful for seismic hazard assessment. Evidence of liquefaction, inundation by tsunami, and seismically induced ground failures may be observed in natural or man-made exposures of young sediments or in sediment cores. If secondary effects disturb the environment, and if they are preserved, then these effects can be cataloged to develop a chronology of paleoearthquakes. The dates of shaking events may be constrained by geochronological methods.

4.21.2.3.4.1 Liquefaction

Liquefaction is a loss of bearing strength that occurs when saturated cohesionless sediments are subjected to strong shaking or cyclical loading. Fluidized sediment may mobilize during liquefaction to form distinctive sedimentary or geomorphic structures such as sand blows, sand dikes, and craters (Obermeier, 1996, 2009). Ground failures such as lateral spreading, slumping, lurching, and cracking may occur in soft sediment. In submarine or lacustrine environments, deposition of turbidites (Obermeier, 2009) or mixed layer clays (Doig, 1998; Marco et al., 1996) can provide evidence of a paleoearthquake. Nonseismic events can create structures that are virtually indistinguishable from seismically deformed sediments, or seismites (Ricci Lucchi, 1995). Therefore, paleoseismologists must correlate candidate seismites over regions and rule out nontectonic origins before concluding that an earthquake occurred. For example, the widespread presence of features attributed to liquefaction and ground failure provides strong evidence of the AD 1700 Cascadia earthquake (Obermeier and Dickinson, 2000) and repeated large intraplate earthquakes in the Mississippi River Valley of the New Madrid Seismic Zone in the central United States (Vidale et al., 2011).

4.21.2.3.4.2 Tsunami, turbidites, and subduction earthquakes

The largest earthquakes recorded in the last century with modern instrumentation have all occurred on subduction zones. Since great earthquakes have long recurrence times, paleoseismology can provide important insights into their frequency, magnitudes, locations, and effects (Goldfinger et al., 2013).

However, most subduction zones are offshore where traditional paleoseismic methods such as fault excavation cannot be applied. Nonetheless, large events leave signatures in the geologic record that can be analyzed to develop a chronology of subduction zone earthquakes. Subduction earthquakes typically involve rapid, large elevation changes that affect the coastal environment. Uplift causes 'kill zones' of marine organisms above sea level, and areas of subsidence lead to drowning of freshwater biota and soils due to inundation by seawater. Vertical changes in seafloor can generate tsunami that inundate low-lying coastal areas and leave distinctive signatures such as widespread scour marks capped by sand. The 2011 M_w 9.0 Tohoku earthquake generated a large destructive tsunami that inundated areas of coastal northeast Japan that contained tsunami deposits from prior large earthquakes (Goldfinger et al., 2013; Tajima et al., 2013). The significance of those tsunami deposits for indicating the great magnitude of past earthquakes was not fully appreciated prior to the 2011 Tohoku event.

The AD 1700 Cascadia earthquake offshore the Pacific Northwest coast of the United States and British Columbia, Canada, left evidence in the form of tsunami deposits, drowned coastal forests, and historical accounts (Adams, 1996; Atwater et al., 1995; Satake et al., 1996). Shaking from the AD 1700 and earlier Cascadia subduction earthquakes also caused offshore sediment gravity flows, leading to the deposition of turbidites. Turbidites are distinctive sedimentary deposits from sediment gravity flows that are triggered by either shaking or storms (Goldfinger et al., 2012). A paleoseismic record of subduction zone earthquakes can be compiled from turbidites revealed in offshore sediment cores.

4.21.2.3.4.3 Landslides and fragile geologic features

Earthquake ground motions often trigger slope failures. Mountainous regions with steep slopes or unstable slope materials are particularly susceptible to seismically induced slope failures. Slope failures have been responsible for many fatalities in historic earthquakes (Bolt, 1999). In seismically active regions, slope failures over large areas may be linked to paleoearthquakes. Several methods have been developed to date paleoearthquakes by measuring the age of seismically induced landslide deposits. For example, lichenometry has been applied to date seismically induced rockfalls during the past several centuries along the Alpine fault zone, New Zealand (Bull, 1996). The criteria for determining seismic origin and methods for dating are summarized by Jibson (1996).

Precariously balanced rocks are fragile geologic features that are susceptible to toppling by strong ground motions. Thus, precariously balanced rocks apparently exist only in places where the ground motions have been relatively low (Anderson et al., 2011). The minimum amplitude acceleration ground motion required to topple a precariously balanced rock can be calculated from field measurements of the rock height and width. The locations and characteristics of precariously balanced rocks provide constraints on maximum ground motions over the time period in which the rocks remained precarious (e.g., Brune, 2002).

4.21.2.3.5 Measurement of slip

Measurement of slip, or surface displacement, across a fault yields information about the magnitude of paleoearthquakes and, when combined with age control, the average rate of deformation. The slip rate of a fault provides an upper bound for the rate of seismic moment (WGCEP, 1995, 2008; Youngs and Coppersmith, 1985). Slip from paleoearthquakes can be used in combination with slip rate to estimate the recurrence intervals between earthquakes and to constrain the magnitude of past ruptures.

4.21.2.3.5.1 Slip per earthquake

The amount of slip per event can be used to estimate the magnitude of previous earthquakes and the time between earthquakes if the average slip rate is known. The distribution of slip may also reveal the location of segment boundaries and areas of repetitive or characteristic slip, if any are present (Ward, 1997).

To measure slip per paleoearthquake, it is necessary to identify a piercing line that has been displaced by a known number of earthquakes. Ideally, this is done by measuring the displacement of a piercing line of known age in an area where the dates of individual surface ruptures are well constrained (e.g., Liu et al., 2004; Liu-Zeng et al., 2006). For normal and reverse faults, the vertical component of displacement can be measured directly by excavating a trench across the scarp (**Figure 11**) because material deposited by the collapse or decay of a vertical scarp typically buries and preserves evidence of the rupture. Changes in slope or height of a scarp along strike may indicate multiple ruptures. Variation in the amount of lateral displacement of piercing lines along a strike-slip fault can indicate multiple surface ruptures. In coastal areas, the vertical component of displacement can be measured from elevated marine shorelines (Matsuda et al., 1978; Plafker and Rubin, 1978). For strike-slip faults, the best measurements of slip per earthquake are obtained from three-dimensional excavation using multiple trenches (Liu-Zeng et al., 2006; Sieh, 1984; Weldon et al., 2004).

An alternate but less reliable method is to measure the displacement of surficial features such as laterally offset channels (**Figure 12**), infer the number of earthquakes that formed each offset, and then derive an inferred slip value (Sieh, 1978; Sieh and Jahns, 1984). Uncertainty in the resulting slip per event measurements is unconstrained. Digital imaging (Bevis et al., 2005; Haddad et al., 2012) and high-resolution DEMs reveal many offset features that could not be recognized or modeled with traditional field mapping and measurement methods (Arrowsmith and Zielke, 2009). These methods can be used to measure offsets from multiple earthquakes (Zielke et al., 2010, 2012). Stratigraphic control is needed to determine the number of earthquakes that generate each offset and thereby measure slip per earthquake (Grant Ludwig et al., 2010).

4.21.2.4 Magnitude of Paleoearthquakes

In general, the effects of an earthquake are proportional to its size (Slemmons and dePolo, 1986). Therefore, the amount of slip, tectonic deformation, or secondary effects preserved in the paleoseismic record helps constrain the size of a paleoseismic event. The magnitude of paleoearthquakes can be estimated from measurements of slip per event or from rupture length inferred by correlating the dates of paleoseismic events at different sites.

Empirical regression relationships and other statistical relationships between historic earthquake magnitude and rupture length, average slip, or maximum slip (Biasi and Weldon, 2006; Wells and Coppersmith, 1994) can be applied to estimate the

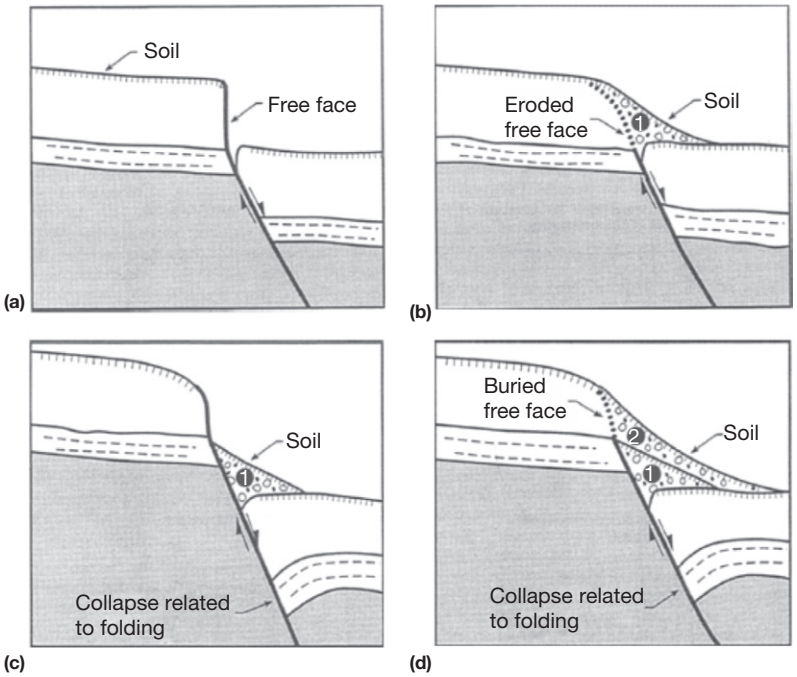


Figure 11 Idealized sequence of two normal faulting events and formation of a colluvial wedge, as exposed in trench excavations. Reproduced from Lettis WR and Kelson KI (2000) Applying geochronology in paleoseismology. In: Noller JS, Sowers JM, and Lettis WR (eds.) *Quaternary Geochronology: Methods and Applications*. *AGU Reference Shelf*, *Series*, vol. 4, pp. 479–495. Washington, DC: American Geophysical Union.

Figure 12 Photograph of a channel offset right laterally ~1 m by faulting along a subsidiary trace of the San Andreas fault in the Carrizo Plain, CA. Reproduced from Grant Ludwig L, Akciz SO, Cheiffetz T, and Sato T (2013) Discovery and paleoseismic investigation of the Bidart fault, a subparallel oblique-slip strand of the San Andreas fault in the Carrizo Plain. *Seismological Research Letters* 84(2): 332.

magnitude of paleoearthquakes (**Table 2**). Hemphill-Haley and Weldon (1999) developed empirical relationships specifically for estimating magnitude from paleoseismic measurements of slip. Their method provides uncertainties based on the number of measurements and percentage of rupture length sampled.

For faults with chronologies of earthquakes from multiple paleoseismic investigations, surface rupture lengths can be estimated by assuming that ruptures with overlapping time windows occurred at the same time (Biasi and Weldon, 2009). However, there are large uncertainties in correlating ruptures between paleoseismic investigation sites, so the inferred rupture length and resulting magnitudes are poorly constrained. Segmentation models (see **Section 4.21.3.1**) have also been applied to estimate paleoearthquake rupture length and magnitude.

4.21.2.5 Slip Rate

The slip rate of a fault, **V**, is obtained by measuring the displacement, **D**, across a fault zone during time interval, **T**, using eqn [1]:

$$\mathbf{V} = \mathbf{D}/\mathbf{T} \qquad [1]$$

Similar to measuring slip, to measure a slip rate, a geologist must find a feature that crosses the fault, is well defined, and is datable. Such features are called *piercing lines*. The points where they intersect a fault zone are called *piercing points*. Fault slip is measured from the displacement of piercing points and deformation of piercing lines. Geologic piercing lines include streams, gullies, ancient shorelines, alluvial fans, and linear sedimentary features. Anthropogenic piercing lines such as walls, orchards, fences, and roads may be useful if the fault slip rate is high and the man-made features have existed for decades to centuries (Rockwell et al., 2002a). The ideal piercing line crosses the fault at a right angle, is perfectly linear, has the same trend on either side of the fault zone, and is of precisely known age. Such conditions rarely exist. Therefore, most slip rates have considerable measurement uncertainty.

Table 2 Selected empirical relationships between moment magnitude (**M**), average displacement (AD), maximum displacement (MD), and surface rupture length (SRL)

Equation[a]	*Slip type*[b]	a *(sa)*[c]	b *(sa)*[c]	*Standard deviation*	*Correlation coefficient*
$\mathbf{M} = a + b^*\log(\mathrm{AD})$	SS	7.04(0.05)	0.89(0.09)	0.28	0.89
	R[d]	6.64(0.16)	0.13(0.36)	0.50	0.10
	N	6.78(0.12)	0.65(0.25)	0.33	0.64
	All	6.93(0.05)	0.82(0.10)	0.39	0.75
$\mathbf{M} = a + b^*\log(\mathrm{MD})$	SS	6.81(0.05)	0.78(0.06)	0.29	0.90
	R[d]	6.52(0.11)	0.44(0.26)	0.52	0.36
	N	6.61(0.09)	0.71(0.15)	0.34	0.80
	All	6.69(0.04)	0.74(0.07)	0.40	0.78
$\mathbf{M} = a + b^*\log(\mathrm{SRL})$	SS	5.16(0.13)	1.12(0.08)	0.28	0.91
	R[d]	5.00(0.22)	1.22(0.16)	0.28	0.88
	N	4.86(0.34)	1.32(0.26)	0.34	0.81
	All	5.08(0.10)	1.16(0.07)	0.28	0.89

[a]Displacement in meters. Surface rupture length in kilometers.
[b]SS, strike-slip; R, reverse; N, normal; All, all fault types.
[c]Coefficients and standard errors.
[d]Relationship is not significant at the 95% confidence level.
Reproduced from Wells DL and Coppersmith KJ (1994) New empirical relationships among magnitude, rupture length, rupture area, and surface displacement. *Bulletin of the Seismological Society of America* 84: 974–1002.

Uncertainties in slip rates are caused by uncertainties in measurements of displacement and in estimates of the time interval over which the displacement occurred (e.g., Behr et al., 2010; Keller et al., 1982; van der Woerd et al., 2006). Aperture of measurement affects slip rates because deformation may occur beyond the main fault zone (McGill and Rubin, 1999; Rockwell et al., 2002b) and have only subtle effects on piercing lines (Grant and Donnellan, 1994; Grant and Sieh, 1993; Sieh, 1996). Therefore, geologically measured slip rates may provide only minimum values if measured over short apertures in regions where deformation over a broader zone cannot be ruled out. For blind faults, the displacement on the fault plane must be inferred from deformation far outside the fault zone, and the resulting slip rates have even larger uncertainties.

The length and variation of time intervals between surface ruptures may affect the measurements of slip rate or slip per event. For faults that do not creep, surface displacements are generated episodically by earthquakes. If the slip rate of a fault is low and the average time between surface ruptures is long – the case for most faults – the measurements of slip rate from features offset in only one or two earthquakes may be significantly greater or less than the average rate over multiple rupture cycles. It is generally assumed that slip rate is constant over the interval of measurement, but for most faults, there are insufficient data to test this assumption. Geomorphic analysis of normal fault scarps in the Basin and Range Province, United States, reveals significant temporal variation in slip rates for several faults (Slemmons, 1995), and irregular earthquake recurrence along the Aksu thrust in Tianshan, China, led to irregular strain release and slip rate (Hubbert-Ferrari et al., 2005).

4.21.3 Models and Interpretation of Data

Paleoseismic data provide information about earthquakes over scales of time and magnitude that are useful for seismic hazard assessment and essential for understanding the long-term rupture patterns of faults. Therefore, paleoseismic data have been influential in the development and testing of models that describe fault behavior over multiple rupture cycles (e.g., Rundle et al., 2006; Smith and Sandwell, 2006) and in forecasting earthquakes (e.g., Rundle et al., 2005; Van Aalsburg et al., 2010). This section describes models that depend critically on observations of paleoearthquakes and discusses uncertainty in paleoseismic data. Several classic models based on inferred patterns of fault slip are shown in **Figure 13**.

Patterns of fault rupture in space and time yield insights about the physics of faulting (Ward, 1997; Weldon et al., 2004, 2005) and provide templates for estimating the location and size of future earthquakes. Several models of fault behavior shown in **Figure 13** (see caption for explanation of diagram) are based on multiple paleoseismic measurements of slip at different sites along a fault. The variable slip model (**Figure 13(a)**) assumes a constant slip rate along the fault with variable displacement per event at a location and variable earthquake size. The uniform slip model (**Figure 13(b)**) is similar to the characteristic slip model (**Figure 13(c)**) because both assume constant-size displacement at a location. The slip rate is constant along the fault in the uniform slip model. Therefore, a range of earthquake sizes is required to account for sections of fault with large or small slip per event. The overlap model assumes nearly constant displacement per event at a point (characteristic slip), variable slip rate along the fault, and repetition of constant-size large earthquakes. Finally, the coupled model is a uniform slip model coupled with a segmentation model. Segmentation and characteristic models are discussed in more detail in the succeeding text.

Spatial and temporal patterns of fault rupture can be analyzed by plotting the dates of paleoearthquakes on a space–time diagram as shown in **Figure 4**. Uncertainty in dates derived from radiocarbon ages can range from decades to

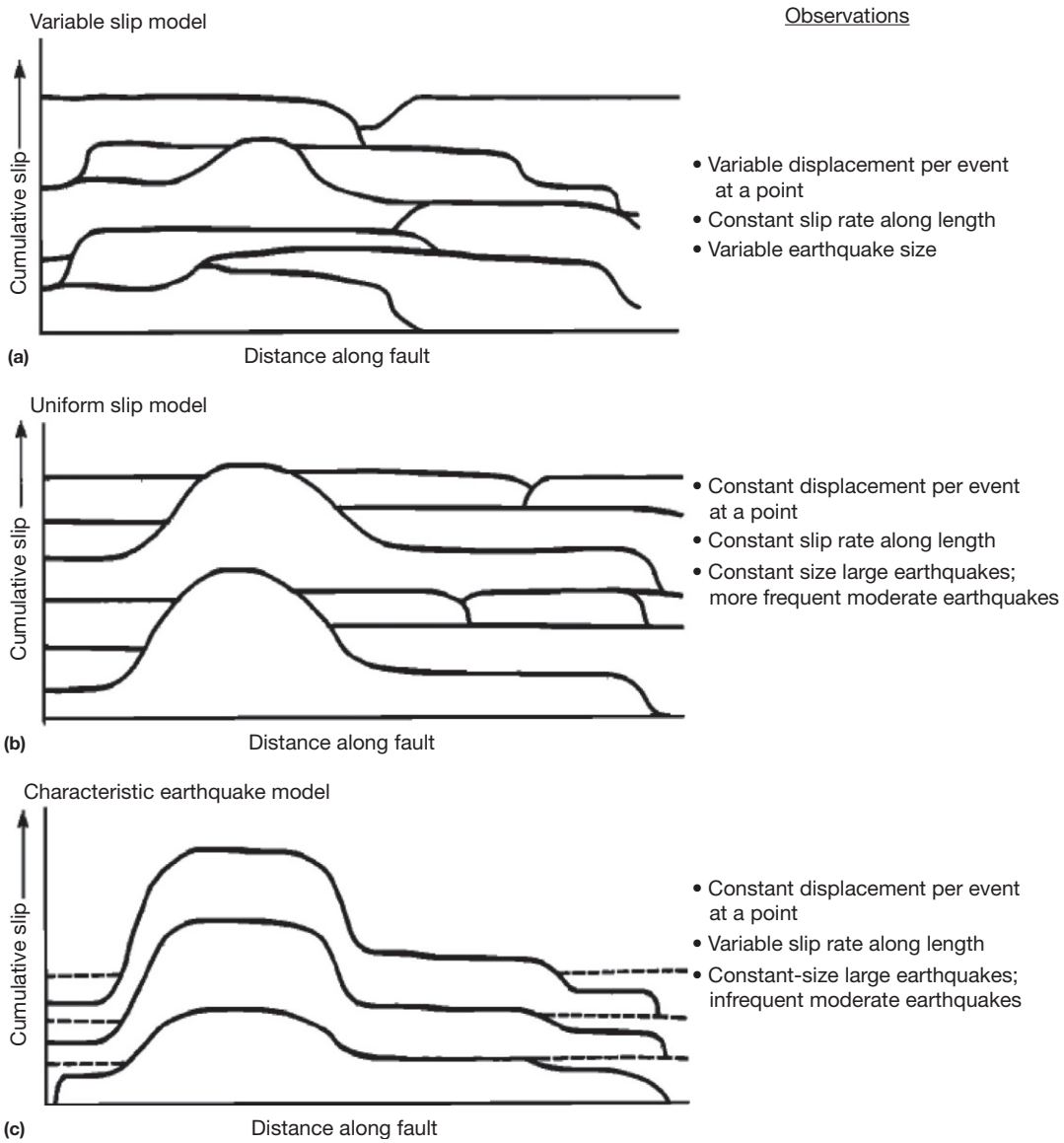

Variable slip model
Observations
Cumulative slip
• Variable displacement per event at a point
• Constant slip rate along length
• Variable earthquake size
(a)
Distance along fault
Uniform slip model
Cumulative slip
• Constant displacement per event at a point
• Constant slip rate along length
• Constant size large earthquakes; more frequent moderate earthquakes
(b)
Distance along fault
Characteristic earthquake model
Cumulative slip
• Constant displacement per event at a point
• Variable slip rate along length
• Constant-size large earthquakes; infrequent moderate earthquakes
(c)
Distance along fault

Figure 13 Illustrations of three models of displacement associated with large-magnitude earthquakes. The cumulative slip distribution is illustrated for each model, and the implications of each model relative to observational data are given. Dashed lines represent ruptures from adjacent segments. Reproduced from Schwartz DP and Coppersmith KJ (1984) Fault behavior and characteristic earthquakes: Examples from the Wasatch and San Andreas fault zones. *Journal of Geophysical Research* 89: 5681–5698.

more than a thousand years. Paleoearthquakes at adjacent sites with overlapping dates could have been the same rupture or separate ruptures that are indistinguishable within the uncertainty. Additional dates of paleoearthquakes would provide better constraints on the locations and magnitudes of past ruptures, and measurements of slip in individual paleoearthquakes would make it easier to assess which events at one site are likely to have been large enough to propagate to other sites (Hemphill-Haley and Weldon, 1999).

4.21.3.1 Segmentation Models

Segmentation models assume that faults are divided into discrete, identifiable sections that behave distinctively over multiple rupture cycles (e.g., Schwartz and Coppersmith, 1984, 1986; Schwartz and Sibson, 1989). Segment boundaries are sections of a fault that might control the termination or initiation of fault ruptures and therefore limit the magnitude and rupture patterns of earthquakes. Fault segments are defined by structural discontinuities such as step-overs that are large enough to terminate ruptures (Wesnousky, 2006), changes in strike, and changes in rheology (Allen, 1968) and by paleoseismic data (Schwartz and Coppersmith, 1984; Schwartz and Sibson, 1989; Sibson, 2002; Sieh and Jahns, 1984; WGCEP, 1988, 2008).

Segmentation models are attractive because they simplify fault behavior. For example, in seismic hazard assessment, it is desirable to estimate the size of the largest earthquake that can

occur on a fault (Mualchin, 2005; Yeats et al., 1997). The maximum magnitude earthquake can be estimated from empirical relationships (**Table 2**) and the estimated length of maximum rupture (Wells and Coppersmith, 1994). This would be the maximum length either of the fault or of the segment. Similarly, the size of a paleoearthquake can be estimated by assuming that the entire segment ruptured.

Testing of segment models requires defining segments and then observing multiple ruptures. This would take decades to centuries for the fastest slipping faults and thousands of years for most faults. Therefore, few segment models have been tested against rupture patterns of historic earthquakes (Sieh, 1996). Results are mixed. For example, two large earthquakes on strike-slip faults in the western United States (M7 1989 Loma Prieta and M7.3 1992 Landers) propagated through previously identified segment or fault boundaries (Sieh et al., 1993; WGCEP, 1990). However, several historic ruptures of normal faults in the Great Basin have terminated at or near fault discontinuities (Zhang et al., 1999). The analysis of paleo-slip data from the San Andreas fault in the Carrizo Plain, CA, from the section of the fault that last ruptured in 1857 ($\sim M$ 7.8), suggests that rupture patterns are not consistent with a segmentation model (Grant Ludwig et al., 2010; Zielke et al., 2010, 2012). More observations are needed to test segmentation hypotheses and their utility for seismic hazard assessment (Scharer, 2010).

4.21.3.2 Characteristic Earthquake Model

The characteristic earthquake model overlaps with segmentation models because it assumes that faults rupture in segments. The basic tenet of the model is that most surface slip on a fault occurs in *characteristic earthquakes*. As originally defined (**Figure 13**), the model 'postulates that individual faults and fault segments tend to generate essentially same size or characteristic earthquakes having a relatively narrow range of magnitudes near the maximum' (Schwartz and Coppersmith, 1984, p. 5681). The magnitude and frequency of occurrence of these maximum earthquakes are provided by paleoseismologic investigations. A premise of the model is that characteristic earthquakes produce *characteristic slip* that repeats at a location. 'The characteristic earthquake model is based on the assumption that the distribution of slip associated with the characteristic event along a fault segment is repeated in successive events...' such that '... the single-event displacements at any point along the fault are essentially constant. However, because the original slip distribution is repeated in successive earthquakes, the slip rate will vary along the length of the fault' (Schwartz and Coppersmith, 1984, p. 5692).

The characteristic slip tenet of the model has important consequences for the shape of the frequency–magnitude distribution of earthquakes on a fault (**Figure 14**). As described by Schwartz and Coppersmith, 'A direct implication of the characteristic earthquake hypothesis is that the occurrence of earthquakes on individual faults and fault segments does not follow a log linear frequency–magnitude relationship of the form ($\log N = a - bM$) described by Gutenberg and Richter (1956)' (p. 5681) and 'the characteristic earthquake model implies a nonlinear earthquake frequency – magnitude relationship that is dominated by the characteristic event and that has a low b value in the moderate-magnitude range' (p. 5692). Because most of the seismic moment is released by repetition of large characteristic earthquakes of approximately the same magnitude, there is a kink in the frequency–magnitude curve at magnitude M' (**Figure 14**), and 'the characteristic earthquake, which is probably also the maximum event, may be more likely to occur than a smaller noncharacteristic earthquake' (Schwartz and Coppersmith, 1984, p. 5696). As shown in **Figure 14**, geologic (i.e., paleoseismic) data are required to constrain the frequency of the maximum magnitude, or characteristic earthquake, that defines the shape of the frequency–magnitude curve.

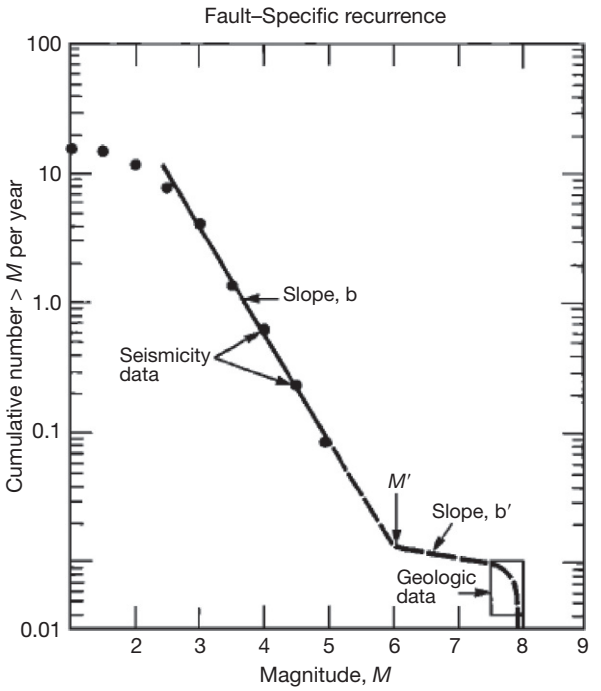


Figure 14 Diagrammatic cumulative frequency–magnitude recurrence relationship for an individual fault or fault segment, as defined by Schwartz and Coppersmith (1984, p. 5696). Large-magnitude characteristic earthquakes are represented by the box labeled 'geologic data.' Historic seismicity is fit with a solid line. Moderate magnitude earthquakes are represented by a dashed line because 'characteristic earthquakes are occurring at the expense of the moderate-magnitude events.' Reproduced from Schwartz DP and Coppersmith KJ (1984) Fault behavior and characteristic earthquakes: Examples from the Wasatch and San Andreas fault zones. *Journal of Geophysical Research* 89: 5681–5698.

Schwartz and Coppersmith (1984) proposed that the characteristic earthquake model helps to define 'fault-specific recurrence' as shown in **Figure 14**, but does not apply worldwide or to regions containing many active faults. Wesnousky (1994) reported that the frequency–magnitude distribution for individual faults is consistent with a characteristic recurrence curve and Kagan (1993) found that a characteristic model does not fit global seismicity rates and an exponential model. The characteristic earthquake model and truncated exponential models are widely used for seismic hazard assessments (Hecker et al., 2013) although there are disparities between observed seismicity and seismicity rates predicted from the characteristic earthquake model (Field et al., 2009).

Rupture forecasts for the San Andreas fault have been based on a combination of segmentation, time-predictable (see succeeding text), and characteristic earthquake models (Field, 2007; Field et al., 2009). The Working Group on California Earthquake Probabilities (WGCEP, 1988, 1990, 1995) divided the San Andreas fault into characteristic segments and issued probabilistic rupture forecasts. In 1989, the M7 Loma Prieta earthquake occurred on portions of two segments defined prior to the earthquake (WGCEP, 1990) rather than on a single segment. The Parkfield section of the San Andreas fault was predicted to break before 1993 with repetition of the 1966 characteristic earthquake (Bakun and Lindh, 1985; WGCEP, 1988). The predicted rupture did not occur until 2004, and although several aspects of the rupture were similar to previous Parkfield earthquakes, other features were unexpected (Harris and Arrowsmith, 2006; Toké et al., 2006).

Several studies have attempted to test the characteristic earthquake model through quantitative analysis of seismicity and fault slip data (e.g., Hecker et al., 2013; Wesnousky, 1986). Hecker et al. (2013) analyzed paleoseismic data for many faults and concluded that many faults exhibit characteristic slip. Characteristic earthquakes have also been reported from Greece (Papadopoulos et al., 2003) and Turkey (Klinger et al., 2003). The San Andreas fault in the Carrizo Plain is one of two keystone areas where data were collected for Schwartz and Coppersmith's (1984) analysis of geologic data, leading to their conclusions about the characteristic slip. More recent findings reveal greater complexity in spatial and temporal rupture history of the San Andreas fault in the Carrizo Plain (Grant, 1996; Grant Ludwig et al., 2010; Grant and Sieh, 1993, 1994). Such complexity in slip was not evident when the characteristic earthquake model was first proposed. These findings challenge the underlying characteristic slip tenet of the model (Grant, 1996; Grant Ludwig et al., 2010; Scharer, 2010). Because the timescale for repetition of 'characteristic earthquakes' is longer than the period of instrumental observations, debate about the applicability of the characteristic earthquake model is likely to continue.

4.21.3.3 Time-Predictable and Recurrence Models

Recurrence time, $\mathbf{T_r}$, is the time interval between successive ruptures of the same fault. $\mathbf{T_r}$ is important for calculations of rupture probability (e.g., Field et al., 2009). Paleoseismic investigations are the main source of data for measuring or estimating $\mathbf{T_r}$. Individual recurrence times or average recurrence intervals (the average time interval between ruptures) can be measured directly by dating successive surface ruptures. For well-studied faults such as the San Andreas, average recurrence intervals and their variance are available for several locations along the fault. For most faults, there is insufficient information to describe variability in recurrence times. If dates of paleoearthquakes are not available, recurrence time may be estimated for a fault using the relationship

$$\mathbf{T_r = D/V} \qquad [2]$$

where $\mathbf{V}$ is the slip rate and $\mathbf{D}$ is the average coseismic surface displacement (slip). Assuming the slip rate is constant over the period of observation and there is no creep, then the recurrence time is a linear function of coseismic displacement. If we assume that displacement per earthquake is constant, then the recurrence time can be 'predicted' by eqn [2]. Thus, eqn [2] is referred to as the time-predictable model, after Shimazaki and Nakata (1980) who developed it from observations of uplifted coastal terraces in Japan. If the constant-displacement-at-a-point assumption and time-predictable model are combined, the result is a periodic recurrence model.

The analysis of historic earthquakes and paleoseismic data from several regions shows that occurrence of large earthquakes is irregular (Goes, 1996). Surface ruptures of the San Andreas fault have been proposed to occur irregularly (Jacoby et al., 1988; Sieh et al., 1989) or in clusters (Grant and Sieh, 1994) or in a quasiperiodic pattern (Scharer et al., 2010). Temporal and/or spatial clustering of surface ruptures has also been proposed to describe surface ruptures in the Basin and Range Province (McCalpin, 1996; Slemmons, 1995; Wallace, 1981) and eastern California shear zone, United States (Rockwell et al., 2002b). Historic ruptures of the North Anatolian fault zone in Turkey (Hartleb et al., 2003, 2006) and patterns of late Holocene seismicity in Iran suggest that seismicity may be triggered by stress shadows from previous ruptures to form temporal clusters or sequences of damaging earthquakes (Berberian and Yeats, 1999; Stein et al., 1996). For most faults, the number of documented paleoearthquakes and the precision of their dates are inadequate to test the hypotheses of clustered or triggered earthquakes (Biasi et al., 2002). Larger paleoseismic data sets are needed to conduct statistically significant tests of recurrence models.

4.21.3.4 Interpretation of Paleoseismic Data

The quality and reliability of paleoseismic data are affected by the investigation method, the characteristics of the study site, and the perspective of the investigator. Many paleoseismic studies are destructive to the study site, so the data cannot be reproduced or compared with a control. Therefore, each paleoseismic publication becomes a repository of data that will be interpreted by scientists who will not be able to evaluate it firsthand. The collection of paleoseismic data requires recognition and interpretation of complex patterns in the geologic record and landscape. Therefore, some interpretation is automatically convolved with the data, and no deconvolution algorithm can be applied after the trench is backfilled. Paleoseismologists have developed several methods to reduce subjectivity and standardize uncertainty in their data. Fault exposures are generally documented with photographs and logs. Photos are objective records, devoid of interpretation though they do not record all of the properties of the sediments that can be observed in the field. 'Trench reviews' are held to allow reviewers to examine primary field data prior to backfill of the trench and publication (**Figure 15**). When possible, care is taken to fill excavations without collapsing the walls so that they can be reopened at a later date. Uncertainty in data collection and field interpretations are often reported qualitatively. For example, Sieh (1978) reported the apparent 'quality' of his measurements of offset streams. Paleoseismologists may report high-quality measurements with 'high confidence.' Such assessments are qualitative and should not

Figure 15 Photograph of trench review conducted in the field prior to trench closure. Professor J. Ramon Arrowsmith of Arizona State University (second from left), Professor Lisa Grant Ludwig (pointing), and Dr. Sinan Akciz (black shirt) of University of California, Irvine, show exposures of sediments in the San Andreas fault zone at Bidart Fan to scientific reviewer, Professor Ray Weldon (far right) of the University of Oregon while an ASU graduate student (far left) watches.

be confused with statistical confidence intervals (Scharer et al., 2007).

4.21.3.4.1 Uncertainty in paleoseismic data and implications for the reliability of seismic hazard assessment

The state of practice for probabilistic seismic hazard assessment and calculation of earthquake probabilities for policymaking in the United States requires the application of the best available science and utilization of the best available data (Petersen et al., 2008; WGCEP, 2008). Determination and evaluation of the best available data is a subjective process where the epistemic uncertainty of expert opinion is represented as weighted branches in logic trees. Paleoseismic data, including slip rate and average recurrence interval, are critical components of seismic hazard calculations but contain inherently large, generally unquantifiable uncertainties due to irreproducibility and interpretive error associated with human judgment. These uncertainties have been difficult to incorporate in seismic hazard analysis even though the epistemic (model-dependent) uncertainty of human judgment may be significant.

Methods for evaluating subjective paleoseismic data have been developed and described by Scharer et al. (2007). These methods involve compilation of trench log data and assignment of qualifiers such as very likely, likely, probable, possible, and very unlikely to evidence that has been subjectively labeled as qualified, satisfactory, sufficient, or somewhat ambiguous in field observations. While this semiformal approach is an improvement over previous practices, it is still subjective and relies on expert opinion to qualify data that are not reproducible and cannot be reviewed by others after trenches are closed. This problem could potentially be addressed by constraining interrater reliability as a proxy for uncertainty associated with expert opinion and judgment in interpreting paleoseismic data (Grant Ludwig and Runnerstrom, 2010). Interrater reliability is commonly used to measure agreement between multiple raters in applications such as evaluations of courtroom testimony, judgment of student performance, and medical diagnoses. It is widely used in the social/behavioral sciences and has many applications in the medical fields and public health. A pilot study (Gould, 2004) suggests that interrater reliability could be applied to quantify epistemic uncertainty in paleoseismic data and thereby improve the reliability of seismic hazard assessments.

Acknowledgments

This chapter was improved by constructive comments of S. F. McGill, S. O. Akciz, and H. Kanamori.

References

Adams J (1996) Paleoseismicity of the Cascadia subduction zone: Evidence from turbidites off the Oregon–Washington margin. *Tectonics* 9: 569–583.

Akciz SO, Grant Ludwig L, Arrowsmith JR, and Zielke O (2010) Century-long average time intervals between earthquake ruptures of the San Andreas fault in the Carrizo Plain, California. *Geology* 38: 787–790.

Akciz SO, Grant LB, and Arrowsmith JR (2009) An average recurrence interval of 144 years for repeated rupture of the San Andreas Fault in the Carrizo Plain, California, since 1280 AD. *Journal of Geophysical Research* 114: B01313.

Allen CR (1968) The tectonic environments of seismically active and inactive areas along the San Andreas fault system. In: Dickinson WR and Grantz A (eds.) Proceedings of Conference on Geologic Problems of the San Andreas Fault System, *Geological Sciences*, vol. 11, pp. 70–82. Stanford, CA: Stanford University Publication.

Anderson JG, Brune JN, Biasi G, Anooshehpoor A, and Purvance M (2011) Workshop report: Applications of precarious rocks and related fragile geological features to U.S. National Hazard Maps. *Seismological Research Letters* 82(3): 431–441.

Arrowsmith JR and Zielke O (2009) Tectonic geomorphology of the San Andreas Fault zone from high resolution topography: An example from the Cholame segment. *Geomorphology* http://dx.doi.org/10.1016/j.geomorph.2009.01.002, special issue on high resolution topography.

Atwater BF, Nelson AR, Clague JJ, et al. (1995) Summary of coastal geologic evidence for past great earthquakes at the Cascadia subduction zone. *Earthquake Spectra* 11: 1–18.

Bakun WH and Lindh AG (1985) The Parkfield, California earthquake prediction experiment. *Science* 229: 619–624.

Balco G, Purvance MD, and Rood DH (2011) Exposure dating of precariously balanced rocks. *Quaternary Geochronology* 6: 295–303.

Behr WM, Rood DH, Fletcher KE, et al. (2010) Uncertainties in slip-rate estimates for the Mission Creek strand of the southern San Andreas fault at Biskra Palms Oasis, southern California. *Geological Society of America Bulletin* 122: 1360–1377. http://dx.doi.org/10.1130/B30020.1.

Bell JW, Brune JN, Liu T, Zreda M, and Yount JC (1998) Dating precariously balanced rocks in seismically active parts of California and Nevada. *Geology* 26(2): 495–498.

Berberian M and Yeats RS (1999) Patterns of historical earthquake rupture in the Iranian plateau. *Bulletin of the Seismological Society of America* 89(1): 120–139.

Bevis M, Hudnut K, Sanchez R, et al. (2005) The B4 project: Scanning the San Andreas and San Jacinto fault zones. *Eos, Transactions of the American Geophysical Union* 86(52), Fall Meet. Suppl. Abstract H34B-01.

Biasi G and Weldon RJ II (1994) Quantitative refinement of calibrated ^{14}C distributions. *Quaternary Research* 41: 1–18.

Biasi G and Weldon RJ II (2006) Estimating surface rupture length and magnitude of paleoearthquakes from point measurements of rupture displacement. *Bulletin of the Seismological Society of America* 96(5): 1612–1623.

Biasi GP and Weldon RJ (2009) San Andreas fault rupture scenarios from multiple paleoseismic records: Stringing pearls. *Bulletin of the Seismological Society of America* 99: 471–498.

Biasi GP, Weldon RJ, Fumal TE, and Seitz GG (2002) Paleoseismic event dating and the conditional probability of large earthquakes on the southern San Andreas fault, California. *Bulletin of the Seismological Society of America* 92: 2761–2781. http://dx.doi.org/10.1785/0120000605.

Bolt BA (1999) *Earthquakes*, 4th edn. New York: W.H. Freeman and Co.

Bonilla MG (1988) Minimum earthquake magnitude associated with coseismic surface faulting. *Bulletin of the Association of Engineering Geologists* XXV: 17–29.

Bonilla MG and Lienkaemper JJ (1990) Visibility of fault strands in exploratory trenches and timing of rupture events. *Geology* 18: 153–156.

Brune JN (2002) Precarious rock constraints on ground motion from historic and recent earthquakes in southern California. *Bulletin of the Seismological Society of America* 92(7): 2602–2611.

Bull WB (1996) Prehistorical earthquakes on the Alpine fault, New Zealand. *Journal of Geophysical Research* 101: 6037–6050.

Bullard TF and Lettis WR (1993) Quaternary fold deformation associated with blind thrusting, Los Angeles Basin, California. *Journal of Geophysical Research* 98: 8349–8369.

Calais E, Freed A, Mattioli G, et al. (2010) Transpressional rupture of an unmapped fault during the 2010 Haiti earthquake. *Nature Geoscience* 3: 794–799. http://dx.doi.org/10.1038/NGEO992.

California Geological Survey (2002) Guidelines for evaluating the hazard of surface fault rupture, California Department of Conservation, Division of Mines and Geology, Note 49. Available online http://www.conservation.ca.gov/cgs/information/publications/cgs_notes/note_49/Documents/note_49.pdf.

Carver GA and McCalpin JP (1996) Paleoseismology of compressional tectonic environments. In: McCalpin JP (ed.) *Paleoseismology*, pp. 183–270. San Diego: Academic Press.

Crone AJ and Omdahl EM (eds.) (1987) Directions in paleoseismology, Proceedings of Conference XXXIX, *U.S. Geological Survey Open-File Report* 87-673: 1–456.

Doig R (1998) 3000-yr Paleoseismological record from the region of the 1988 Saguenay, Quebec, earthquake. *Bulletin of the Seismological Society of America* 88(5): 1198–1203.

Dolan JF, Stevens D, and Rockwell TK (2000) Paleoseismologic evidence for an early to mid-Holocene age of the most recent surface rupture on the Hollywood fault, Los Angeles, California. *Bulletin of the Seismological Society of America* 90: 334–344.

Field EH (2007) A summary of previous working groups on California earthquake probabilities. *Bulletin of the Seismological Society of America* 97: 1033–1053. http://dx.doi.org/10.1785/0120060048.

Field EH, Dawson TE, Felzer KR, et al. (2009) Uniform California Earthquake Rupture Forecast, Version 2 (UCERF2). *Bulletin of the Seismological Society of America* 99(4): 2053–2107. http://dx.doi.org/10.1785/0120080049.

Fletcher KEK, Sharp WD, Kendrick KJ, Behr WM, Hudnut KW, and Hanks TC (2010) Th-230/U dating of a later Pleistocene alluvial fan along the southern San Andreas fault. *Geological Society of America Bulletin* 122(9–10): 1347–1359. http://dx.doi.org/10.1130/B30018.1.

Goes SDB (1996) Irregular recurrence of large earthquakes: An analysis of historic and paleoseismic catalogs. *Journal of Geophysical Research* 101: 5739–5749.

Goldfinger C, Ikeda Y, Yeats RS, and Ren J (2013) Superquakes and supercycles. *Seismological Research Letters* 84(1): 1–9. http://dx.doi.org/10.1785/0220110135.

Goldfinger C, Nelson CH, Morey A, et al. (2012) Turbidite event history: Methods and implications for Holocene Paleoseismicity of the Cascadia Subduction Zone. In: *U. S. Geological Survey Professional Paper 1661-F,* 184 pp.

Gould MM (2004) *A Method for Populating Paleoseismic Databases using Non-expert Readers*. M.A. Thesis in Social Ecology, University of California, Irvine, 51 pp.

Grant LB (1996) Uncharacteristic earthquakes on the San Andreas fault. *Science* 272: 826–827.

Grant LB (2002) Paleoseismology. In: Lee WH, Kanamori H, and Jennings PC (eds.) *IASPEI International Handbook of Earthquake and Engineering Seismology*, vol. 81A, pp. 475–489. International Association of Seismology and Physics of the Earth's Interior, ch. 30.

Grant L (2007) Historical Seismicity – Paleoseismology. Schubert G (ed.) *Treatise on Geophysics*. In: Kanamori H (ed.) *Earthquake Seismology*, vol. 4, pp. 567–590. Oxford: Elsevier Ltd., ch. 19.

Grant Ludwig L, Akciz SO, Cheiffetz T, and Sato T (2013) Discovery and paleoseismic investigation of the Bidart fault, a subparallel oblique-slip strand of the San Andreas fault in the Carrizo Plain. *Seismological Research Letters* 84(2): 332.

Grant Ludwig L, Akciz SO, Noriega GR, Zielke O, and Arrowsmith JR (2010) Climate-modulated channel incision and rupture history of the San Andreas fault in the Carrizo Plain. *Science* 327(5969): 1117–1119.

Grant Ludwig L and Runnerstrom MG (2010) A first attempt to constrain reliability and measurement error associated with expert opinion and judgment in interpreting paleoseismic data. *Seismological Research Letters* 81(2): 330.

Grant LB, Ballenger LJ, and Runnerstrom EE (2002) Coastal uplift of the San Joaquin Hills, Southern Los Angeles basin, California, by a large earthquake since 1635 A.D. *Bulletin of the Seismological Society of America* 92: 590–599.

Grant LB and Donnellan A (1994) 1855 and 1991 surveys of the San Andreas fault: Implications for fault mechanics. *Bulletin of the Seismological Society of America* 84: 241–246.

Grant LB and Gould MM (2004) Assimilation of paleoseismic data for earthquake simulation. *Pure and Applied Geophysics* 161(11/12): 2295–2306.

Grant LB, Gould MM, Donnellan A, et al. (2005) A web-service based universal approach to heterogeneous fault databases. *Computing in Science and Engineering* 7: 51–57.

Grant LB and Lettis WR (eds.) (2002) Paleoseismology of the San Andreas Fault System. *Bulletin of the Seismological Society of America* 92(7): 2551–2554, special issue.

Grant LB, Mueller KJ, Gath EM, et al. (1999) Late Quaternary uplift and earthquake potential of the San Joaquin Hills, southern Los Angeles basin, California. *Geology* 27: 1031–1034.

Grant LB and Sieh K (1993) Stratigraphic evidence for 7 meters of dextral slip on the San Andreas fault during the great 1857 earthquake in the Carrizo Plain. *Bulletin of the Seismological Society of America* 83: 619–635.

Grant LB and Sieh K (1994) Paleoseismic evidence of clustered earthquakes on the San Andreas fault in the Carrizo Plain, California. *Journal of Geophysical Research* 99(B4): 6819–6841.

Grant LB, Waggoner JT, von Stein C, and Rockwell TK (1997) Paleoseismicity of the North Branch of the Newport-Inglewood fault zone in Huntington Beach, California, from Cone Penetrometer Test Data. *Bulletin of the Seismological Society of America* 87: 277–293.

Gutenberg B and Richter CF (1956) Earthquake magnitude, intensity, energy and acceleration. *Bulletin of the Seismological Society of America* 46: 105–145.

Haddad DE, Akciz SO, Arrowsmith JR, et al. (2012) Applications of airborne and terrestrial laser scanning to paleoseismology. *Geosphere* 8(4): 771–786. http://dx.doi.org/10.1130/GES00701.1, themed issue on seeing the true shape of earth's surface: applications of airborne and Terrestrial Lidar (Frankel volume).

Harris RA and Arrowsmith JR (2006) Introduction to the special issue on the 2004 Parkfield Earthquake and the Parkfield earthquake prediction experiment. *Bulletin of the Seismological Society of America* 96(4B): S1–S10. http://dx.doi.org/10.1785/0120050831, special issue on the 2004 Parkfield Earthquake and Parkfield Experiment.

Hartleb RD, Dolan JF, Akyuz HS, and Yerli B (2003) A 2000-year-long paleoseismologic record of earthquakes along the central North Anatolian fault, from trenches at Alayurt, Turkey. *Bulletin of the Seismological Society of America* 93(5): 1935–1954.

Hartleb RD, Dolan JF, Kozaci O, Akyuz HS, and Seitz GG (2006) A 2500-yr-long paleoseismologic record of large, infrequent earthquakes on the North Anatolian fault and Cukurcimen, Turkey. *Geological Society of America Bulletin* 118: 823–840.

Hecker S, Abrahamson NA, and Woodell KE (2013) Variability of displacement at a point: Implications for earthquake-size distribution and rupture hazard on faults. *Bulletin of the Seismological Society of America* 103(2A): 651–674. http://dx.doi.org/10.1785/0120120159.

Hemphill-Haley MA and Weldon RJ III (1999) Estimating prehistoric earthquake magnitude from point measurements of surface rupture. *Bulletin of the Seismological Society of America* 89(5): 1264–1279.

Hubbert-Ferrari A, Suppe J, Wang X, and Lu HF (2005) Irregular earthquake cycle along the southern Tianshan front, Aksu area, China. *Journal of Geophysical Research - Solid Earth* 110(B6): B06402.

Jacoby G, Sheppard P, and Sieh K (1988) Irregular recurrence of large earthquakes along the San Andreas fault – Evidence from trees. *Science* 241: 196–199.

Jibson RW (1996) Using landslides for paleoseismic analysis. In: McCalpin JP (ed.) *Paleoseismology*, pp. 397–438. San Diego: Academic Press.

Jones LM, Bernknopf R, Cox D, et al. (2008) *The ShakeOut Scenario*, U.S. Geological Survey Open-File Report 2008-1150 and California Geological Survey Preliminary Report 25, http://pubs.usgs.gov/of/2008/1150/.

Kagan Y (1993) Statistics of characteristic earthquakes. *Bulletin of the Seismological Society of America* 83: 7–24.

Keller EA, Bonkowski MS, Korsch RJ, and Shlemon RJ (1982) Tectonic geomorphology of the San Andreas fault zone in the southern Indio Hills, Coachella Valley, California. *Geological Society of America Bulletin* 93(1): 46–56.

Keller EA and Pinter N (1996) *Active Tectonics: Earthquakes, Uplift, and Landscape*. Upper Saddle River, NJ: Prentice Hall.

Klinger Y, Sieh K, Altunel E, et al. (2003) Paleoseismic evidence of characteristic slip on the western segment of the North Anatolian fault, Turkey. *Bulletin of the Seismological Society of America* 93: 2317–2332.

Lajoie KR (1986) Coastal tectonics. In: Wallace RE Chairman (ed.) *Active Tectonics: Studies in Geophysics*, pp. 95–124. Washington, DC: Natl. Acad. Press.

Leon LA, Dolan JF, Shaw JH, and Pratt TL (2009) Evidence for large Holocene earthquakes on the Compton thrust fault, Los Angeles, California. *Journal of Geophysical Research* 114: B12305. http://dx.doi.org/10.1029/2008JB006129.

Lettis WR and Kelson KI (2000) Applying geochronology in paleoseismology. In: Noller JS, Sowers JM, and Lettis WR (eds.) *Quaternary Geochronology: Methods and Applications*. *AGU Reference Shelf, Series*, vol. 4, pp. 479–495. Washington, DC: American Geophysical Union.

Lettis WR, Wells DL, and Baldwin JN (1997) Empirical observations regarding reverse earthquakes, blind thrust faults, and Quaternary deformation: Are blind thrust faults truly blind? *Bulletin of the Seismological Society of America* 87: 1171–1198.

Lienkaemper JJ and Bronk Ramsey C (2009) OxCal: Versatile tool for developing paleoearthquake chronologies – A primer. *Seismological Research Letters* 80(3): 431–434.

Liu J, Klinger Y, Sieh K, and Rubin C (2004) Six similar sequential ruptures of the San Andreas fault, Carrizo Plain, California. *Geology* 32: 649–652.

Liu-Zeng J, Klinger Y, Sieh K, Rubin C, and Seitz G (2006) Serial ruptures of the San Andreas fault, Carrizo Plain, California, revealed by three-dimensional excavations. *Journal of Geophysical Research* 111: B02306. http://dx.doi.org/10.1029/2004JB003601.

Marco S, Stein M, Agnon A, and Ron H (1996) Long term earthquake clustering: A 50,000 year paleoseismic record in the Dead Sea graben. *Journal of Geophysical Research* 101: 6179–6191.

Matsuda T, Ota Y, Ando M, and Yonekura N (1978) Fault mechanism and recurrence time of major earthquakes in southern Kanto district, Japan, as deduced from coastal terrace data. *Geological Society of America Bulletin* 89: 1610–1618.

McCalpin JP (ed.) (1996) *Paleoseismology*. San Diego: Academic Press.

McCalpin JP (ed.) (2009) In: International Geophysics Series, vol. 95. 2nd edn. *Paleoseismology*. *San Diego: Elsevier/Academic Press*.

McCalpin JP and Nelson AR (1996) Introduction to paleoseismology. In: McCalpin JP (ed.) *Paleoseismology*, pp. 1–32. San Diego: Academic Press.

McGill SF and Rubin CM (1999) Surficial slip distribution on the central Emerson fault during the June 28, 1992, Landers earthquake, California. *Journal of Geophysical Research* 104(B3): 4811–4833.

Meltzner AJ, Sieh K, Chiang H-W, et al. (2010) Coral evidence for earthquake recurrence and an A.D. 1390–1455 cluster at the south end of the 2004 Ache-Andaman rupture. *Journal of Geophysical Research – Solid Earth* 115: B10402. http://dx.doi.org/10.1029/2010JB007499.

Mualchin L (2005) Seismic hazard analysis for critical infrastructure in California. *Engineering Geology* 79(3–4): 177–184.

Natawidjaja DH, Sieh K, Chlieh M, et al. (2006) Source parameters of the great Sumatran megathrust earthquakes of 1797 and 1833 inferred from coral microatolls. *Journal of Geophysical Research – Solid Earth* 111(B6): B06403.

Natawidjaja DH, Sieh K, Ward SN, et al. (2004) Paleogeodetic records of seismic and aseismic subduction from central Sumatran microatolls, Indonesia. *Journal of Geophysical Research – Solid Earth* 109(B4): B04306.

Nelson AR, Atwater BF, Bobrowsky PT, et al. (1995) Radiocarbon evidence for extensive plate-boundary rupture about 300 years ago at the Cascadia subduction zone. *Nature* 378: 371–374.

Noller JS, Sowers JM, Colman SM, and Pierce KL (2000a) Introduction to quaternary geochronology. In: Noller JS, Sowers JM, and Lettis WR (eds.) *Quaternary Geochronology: Methods and Applications*. *AGU Reference Shelf*, vol. 4, pp. 1–10. Washington, DC: American Geophysical Union.

Noller JS, Sowers JM, and Lettis WR (2000b) In: *Quaternary Geochronology: Methods and Applications*. *AGU Reference Shelf*, vol. 4. Washington, DC: American Geophysical Union.

Nur A (2007) Historical seismicity – Archeoseismology. In: Kanamori H and Schubert G (eds.) *Treatise on Geophysics*. *Earthquake Seismology*, vol. 4, pp. 591–606. Oxford: Elsevier Ltd., ch. 20.

Obermeier SF (1996) Using liquefaction-induced features for paleoseismic analysis. In: McCalpin JP (ed.) *Paleoseismology*, pp. 331–396. San Diego: Academic Press.

Obermeier SF (2009) Using liquefaction-induced and other soft-sediment features for paleoseismic analysis. In: McCalpin JP (ed.) *Paleoseismology*, pp. 497–564. Elsevier: Academic Press.

Obermeier SF and Dickinson SE (2000) Liquefaction evidence for the strength of ground motions resulting from late Holocene Cascadia subduction earthquakes, with emphasis on the event of 1700 A.D. *Bulletin of the Seismological Society of America* 90: 876–896.

Pantosti D and Yeats RS (1993) Paleoseismology of great earthquakes of the late Holocene. *Annali di Geofisica* 36(3–4): 237–257.

Papadopoulos GA, Karastathis VK, Ganas A, Pavlides S, Fokaefs A, and Orfanogiannaki K (2003) The Lefkada, Ionian Sea (Greece), shock (Mw6.2) of 14 August 2003: Evidence for the characteristic earthquake from seismicity and ground failures. *Earth, Planets and Space* 55(11): 713–718.

Pavlides SB, Pantosti D, and Zhang P (1999) Earthquakes, paleoseismology and active tectonics. *Tectonophysics* 308: 1–2, special issue, vii–x.

Petersen MD, Frankel AD, Harmsen SC, et al. (2008) *Documentation for the 2008 Update of the United States National Seismic Hazard Maps*, U.S. Geological Survey Open-File Report 2008-1128, 61 pp.

Pirazzoli PA (1991) *World Atlas of Holocene Sea-Level Changes*. *Oceanography Series*, vol. 58. Amsterdam: Elsevier.

Plafker G and Rubin M (1978) Uplift history and earthquake recurrence as deduced from marine terraces on Middleton Island, Alaska. *U.S. Geological Survey Open-File Report* 78-943: 687–722.

Prentice CS, Mann P, Crone AJ, et al. (2010) SEismic hazard of the Enriquillo-Plantain Garden fault in Haiti inferred from palaeoseismology. *Nature Geoscience* 3: 789–793. http://dx.doi.org/10.1038/NGEO991.

Prentice CS, Schwartz DP, Yeats RS, and Conveners (1994) Proceedings of the Workshop on Paleoseismology, 18–22 September 1994, Marshall, California, *U.S. Geological Survey Open-File Report* 94-568: 1–210.

Ramsey CB and Lee S (2013) Recent and planned developments of the program OxCal. *Radiocarbon* 55(2–3): 720–730.

Reiter L (1991) *Earthquake Hazard Analysis: Issues and Insights*. New York: Columbia University Press.

Reiter L (1995) Paleoseismology – A user's perspective. In: Serva L and Slemmons DB (eds.) *Perspectives in Paleoseismology*, *Association of Engineering Geologists Special Publication*, vol. 6, pp. 3–6.

Ricci Lucchi F (1995) Sedimentological indicators of paleoseismicity. In: Serva L and Slemmons DB (eds.) *Perspectives in Paleoseismology*, *Association of Engineering Geologists Special Publication*, vol. 6, pp. 7–17.

Rockwell TK, Lindvall S, Dawson T, Langridge R, and Lettis W (2002a) Lateral offsets on surveyed cultural features resulting from the 1999 Izmit and Duzce earthquakes, Turkey. *Bulletin of the Seismological Society of America* 90: 1200–1236.

Rockwell TK, Lindvall S, Herzberg M, Murbach D, Dawson T, and Berger G (2002b) Paleoseismology of the Johnson Valley, Kickapoo and Homestead Valley faults: Clustering of earthquakes in the eastern California shear zone. *Bulletin of the Seismological Society of America* 90: 1200–1236.

Rundle JB, Rundle PB, Donnellan A, et al. (2006) Stress transfer in earthquakes and forecasting: Inferences from numerical simulations. *Tectonophysics* 413: 109–125.

Rundle JB, Rundle PB, Donnellan A, et al. (2005) A simulation-based approach to forecasting the next great San Francisco earthquake. *Proceedings of the National Academy of Sciences* 102: 1–5, www.pnas.org/cgi/doi/10.1073/pnas.0507528102.

Satake K, Shimazaki K, Tsuji Y, and Ueda Y (1996) Time and site of a giant earthquake in Cascadia inferred from Japanese tsunami records of January 1700. *Nature* 379: 246–249.

Scharer KM (2010) Changing views of the San Andreas fault. *Science* 327: 1089.

Scharer KM, Biasi GP, and Weldon RJ (2011) A re-evaluation of the Pallett Creek paleoearthquake chronology based on new AMS radiocarbon dates, San Andreas fault, California. *Journal of Geophysical Research* 116: B12111. http://dx.doi.org/10.1029/2010JB008099.

Scharer KM, Biasi GP, Weldon RJ, and Fumal TE (2010) Quasi-periodic recurrence of large earthquakes on the southern San Andreas fault. *Geology* 38: 555–558.

Scharer KM, Weldon RJ II, Fumal TE, and Biasi GP (2007) Paleoearthquakes on the southern San Andreas fault, Wrightwood, California, 3000 to 1500 B.C.: A new method for evaluating paleoseismic evidence and earthquake horizons. *Bulletin of the Seismological Society of America* 97(4): 1054–1093. http://dx.doi.org/10.1785/0120060137.

Schumm SA, Dumont JF, and Holbrook JM (2000) *Active Tectonics and Alluvial Rivers*. Cambridge: Cambridge University Press.

Schwartz DP and Coppersmith KJ (1984) Fault behavior and characteristic earthquakes: Examples from the Wasatch and San Andreas fault zones. *Journal of Geophysical Research* 89: 5681–5698.

Schwartz DP and Coppersmith KJ (1986) Seismic hazards: New trends in analysis using geologic data. In: Wallace RE (ed.) *Active Tectonics: Studies in Geophysics*, pp. 215–230. Washington, DC: National Academy Press.

Schwartz DP and Sibson RH (eds.) (1989) Fault segmentation and controls of rupture initiation and termination. In: *U.S. Geological Survey Open-File Report*, 89-315: pp.1–447.

Serva L and Slemmons DB (eds.) (1995) *Perspectives in Paleoseismology*: Dallas, TX: Association of Engineering Geologists, special publication 6.

Shimazaki K and Nakata T (1980) Time-predictable recurrence model for large earthquakes. *Geophysical Research Letters* 7: 279–282.

Sibson RH (2002) Geology of the crustal earthquake source. In: Lee WHK, Kanamori H, and Jennings PC (eds.) *Handbook of Earthquake and Engineering Seismology*. San Diego: Academic Press.

Sieh KE (1978) Slip along the San Andreas fault associated with the great 1857 earthquake. *Bulletin of the Seismological Society of America* 68: 1421–1448.

Sieh KE (1981) A review of geological evidence for recurrence times for large earthquakes. In: Simpson DW and Richards PG (eds.) *Earthquake Prediction: An International Review*. *Maurice Ewing Series*, vol. 4, pp. 209–216. Washington, DC: American Geophysical Union.

Sieh KE (1984) Lateral offsets and revised dates of large earthquakes at Pallett Creek, California. *Journal of Geophysical Research* 89: 7641–7670.

Sieh K (1996) The repetition of large-earthquake ruptures. *Proceedings of the National Academy of Sciences* 93: 3764–3771.

Sieh KE and Jahns RH (1984) Holocene activity of the San Andreas fault at Wallace Creek, California. *Geological Society of America Bulletin* 95: 883–896.

Sieh KE, Jones L, Hauksson E, et al. (1993) Near-field investigations of the Landers earthquake sequence, April to July 1992. *Science* 260: 171–176.

Sieh KE, Stuiver M, and Brillinger D (1989) A more precise chronology of earthquakes produced by the San Andreas fault in Southern California. *Journal of Geophysical Research* 94: 603–623.

Slemmons DB (1995) Complications in making paleoseismic evaluations in the Basin and Range province, western United States. In: Serva L and Slemmons DB (eds.) Perspectives in Paleoseismology. *Bulletin of the Association of Engineering Geologists* 19–34, special publication 6.

Slemmons DB and dePolo CM (1986) Evaluation of active faulting and related hazards. In: Wallace RE (ed.) *Active Tectonics: Studies in Geophysics*, pp. 45–62. Washington, DC: National Academy Press.

Smith BR and Sandwell DT (2006) A model of the earthquake cycle along the San Andreas fault system for the past 1000 years. *Journal of Geophysical Research – Solid Earth* 111(B1): B01405.

Somerville P and Moriwaki Y (2002) Seismic hazards, risk assessment, and building codes. In: Lee WHK, Kanamori H, and Jennings PC (eds.) *IASPEI International Handbook of Earthquake and Engineering Seismology*. San Diego: Academic Press.

Stein RS, Dieterich JH, and Barka AA (1996) Role of stress triggering in earthquake migration on the North Anatolian fault. *Physics and Chemistry of the Earth* 21: 225–230.

Stein RS and Yeats RS (1989) Hidden earthquakes. *Scientific American* 260(6): 48–57.

Tajima F, Mori J, and Kennett BLN (2013) A review of the 2011 Tohoku-Oki earthquake (Mw 9.0): Large-scale rupture across heterogeneous plate coupling. *Tectonophysics* 586: 15–34. http://dx.doi.org/10.1016/j.tecto.2012.09.014.

Toké NA, Arrowsmith JR, Young JJ, and Crosby C (2006) Paleoseismic and postseismic observations of surface slip along the Parkfield segment of the San Andreas Fault. *Bulletin of the Seismological Society of America* 96(4B): S221–S238. http://dx.doi.org/10.1785/0120050809, special issue on the 2004 Parkfield Earthquake and Parkfield Experiment.

Trumbore SE (2000) Radiocarbon geochronology. In: Noller JS, Sowers JM, and Lettis WR (eds.) *Quaternary Geochronology: Methods and Applications*. *AGU Reference Shelf Series*, vol. 4, pp. 41–60. Washington, DC: American Geophysical Union.

Van Aalsburg J, Rundle JB, Grant L, et al. (2010) Space- and time-dependent probabilities for earthquake fault systems from numerical simulations: Feasibility study and first results. *Pure and Applied Geophysics* 167: 967–977.

Van der Woerd J, Klinger Y, Sieh K, Tapponier P, Ryerson FJ, and Meriaux AS (2006) Long-term slip rate of the southern San Andreas Fault from Be-10-Al-26 surface exposure dating of an offset alluvial fan. *Journal of Geophysical Research - Solid Earth* 111(B4): B044407.

Vidale J, Atkinson G, Green R, et al. (2011) Report of the Independent Expert Panel on New Madrid Seismic Zone Earthquake Hazards to the National Earthquake Prediction Evaluation Council (NEPEC) and Dr. Marcia McNutt, Director of the U.S. Geological Survey, April 16, 2011, http://earthquake.usgs.gov/aboutus/nepec/reports/index.php.

Vittori E, Labini SS, and Serva L (1991) Palaeoseismology; review of the state-of-the-art. *Tectonophysics* 193: 9–32.

Wallace RE (1981) Active faults, paleoseismology, and earthquake hazards in the western United States. In: Simpson DW and Richards PG (eds.) *Earthquake Prediction: An International Review*. *Maurice Ewing Series,* vol. 4, pp. 209–216. Washington, DC: American Geophysical Union.

Wallace RE (1986) *Active Tectonics: Studies in Geophysics*. Washington, DC: National Academic Press.

Ward SN (1997) Dogtails versus rainbows: Synthetic earthquake rupture models as an aid in interpreting geological data. *Bulletin of the Seismological Society of America* 87: 1422–1441.

Weldon RJ, Fumal TE, Biasi GP, and Scharer KM (2005) Past and future earthquakes on the San Andreas fault. *Science* 308: 966–967.

Weldon R, Scharer K, Fumal T, and Biasi G (2004) Wrightwood and the earthquake cycle: What a long recurrence record tells us about how faults work. *GSA Today* 14: 4–10.

Wells DL and Coppersmith KJ (1994) New empirical relationships among magnitude, rupture length, rupture area, and surface displacement. *Bulletin of the Seismological Society of America* 84: 974–1002.

Wesnousky SG (1986) Earthquakes, Quaternary faults, and seismic hazard in California. *Journal of Geophysical Research* 91: 12587–12631.

Wesnousky SG (1994) Gutenberg-Richter or characteristic earthquake distribution: Which one is it? *Bulletin of the Seismological Society of America* 84: 1940–1959.

Wesnousky SG (2006) Predicting the endpoints of earthquake ruptures. *Nature* 444: 358–360.

Working Group on California Earthquake Probabilities (2008) *The Uniform California Earthquake Rupture Forecast, Version 2 (UCERF 2),* Open-File Report 2007-1437 and California Geological Survey Special Report 203, Washington, DC: U.S. Geological Survey.

Working Group on California Earthquake Probabilities (WGCEP) (1988) Probabilities of large earthquakes occurring in California on the San Andreas fault. *U.S. Geological Survey Open-File Report* 88-398: 1–62.

Working Group on California Earthquake Probabilities (WGCEP) (1990) Probabilities of large earthquakes in the San Francisco Bay region, California. *U.S. Geological Survey Circular* 1053: 1–51.

Working Group on California Earthquake Probabilities (WGCEP) (1995) Seismic hazards in southern California – Probable earthquakes, 1994–2024. *Bulletin of the Seismological Society of America* 85: 379–525.

Yeats RS and Prentice CS (1996) Introduction to special section: Paleoseismology. *Journal of Geophysical Research* 101(B3): 5847–5853.

Yeats RS, Sieh KE, and Allen CR (1997) *The Geology of Earthquakes*. New York: Oxford University Press.

Youngs RR and Coppersmith KJ (1985) Implications of fault slip rates and earthquake recurrence models to probabilistic seismic hazard estimates. *Bulletin of the Seismological Society of America* 75: 939–964.

Zachariasen J, Sieh K, Taylor FW, Edwards RL, and Hantoro WS (1999) Submergence and uplift associated with the giant 1833 Sumatran subduction earthquake: Evidence from coral microatolls. *Journal of Geophysical Research - Solid Earth* 104(B1): 895–919.

Zachariasen J, Sieh K, Taylor FW, and Hantoro WS (2000) Modern vertical deformation above the Sumatra subduction zone: Paleogeodetic insights from coral microatolls. *Bulletin of the Seismological Society of America* 90: 897–913.

Zhang PZ, Mao FY, and Slemmons DB (1999) Rupture terminations and size of segment boundaries from historical earthquake ruptures in the Basin and Range Province. *Tectonophysics* 308: 37–52.

Zielke O, Arrowsmith JR, Grant Ludwig L, and Akciz SO (2010) Slip in the 1857 and earlier large earthquakes Along the Carrizo Plain, San Andreas fault. *Science* 327(5969): 1119–1122.

Zielke O, Arrowsmith JR, Grant Ludwig L, and Akciz SO (2012) High resolution topography-derived offsets along the 1857 Fort Tejon earthquake rupture trace, San Andreas Fault. *Bulletin of the Seismological Society of America* 102(3): 1135–1154. http://dx.doi.org/10.1785/0120110230.

Ziony JI and Yerkes RF (1985) Evaluating earthquake and surface faulting potential. In: Ziony JI (ed.) *Earthquake Hazards in the Los Angeles Region*, *U. S. Geological Survey Professional Paper* 1360, pp. 43–91.

This page intentionally left blank

4.22 Historical Seismicity: Archaeoseismology

A Nur, Stanford University, Stanford, CA, USA

© 2007 Elsevier B.V. All rights reserved.

This chapter is reproduced from the previous edition, Volume 4, pp. 591–606, © 2007, Elsevier B.V.

4.22.1	**Overview**	581
4.22.1.1	Frequency–Magnitude Relation and the Historical Observation Window	581
4.22.1.2	Limited Usefulness of History	582
4.22.1.3	Effects of Tectonics on Sites of Human Occupation, Resonant Effects, and Vulnerability of Tells	583
4.22.1.4	Communication Issues Between Disciplines	584
4.22.2	**Earthquake Traces in Archaeological Sites**	585
4.22.2.1	Effects of Earthquakes on Landforms	585
4.22.2.1.1	Fault displacement	585
4.22.2.1.2	Elevated or submerged coastal land surfaces	585
4.22.2.1.3	Tilting of horizontal surfaces	585
4.22.2.2	Architectural Damage	586
4.22.2.2.1	Collapsed walls with preferential axes of destruction	586
4.22.2.2.2	Aligned fallen columns	586
4.22.2.2.3	Slipped keystones	587
4.22.2.3	Skeletons of Trapped Humans and Domestic Animals	588
4.22.2.4	Fire	589
4.22.3	**Techniques in Archaeoseismology**	589
4.22.3.1	Informing Archaeological Methods with a Seismology Perspective	589
4.22.3.2	Intersection of Archaeology and Paleoseismology	590
4.22.3.3	Forensic Anthropology	590
4.22.3.4	Numerical Simulation	590
4.22.4	**Mutual Benefits of Archaeoseismology**	591
4.22.4.1	Benefits to Seismology	591
4.22.4.2	Benefits to Archaeology	591
References		591

4.22.1 Overview

4.22.1.1 Frequency–Magnitude Relation and the Historical Observation Window

Earthquakes recur. One of the fundamental observations of modern earthquake science is that earthquakes occur on faults and that the seismicity on any fault follows a pattern of recurrence over time, called the frequency–magnitude relation.

If we sort the many small earthquakes recorded by seismographs for a given fault, including ones too small to be felt without sensitive instruments, we see a pattern emerge. There are many more small earthquakes than large ones. In fact, a magnitude-3 earthquake is ten times more likely to occur than a magnitude-4 earthquake, which is ten times more likely than a magnitude-5 earthquake, and so on. At first glance, this seems to imply that very large earthquakes do not occur. On the contrary, the distribution actually gives us a specific empirical relationship between earthquake magnitude and recurrence interval on a particular fault, for all earthquake magnitudes, up to the largest that can occur on that fault. This relationship is called the frequency–magnitude relation. If ten magnitude-6 earthquakes occurred in 100 years of recorded seismicity, and one magnitude 7 occurred, we would expect 0.1 magnitude-8 earthquakes to occur in that same period, provided that the fault is long enough to generate an earthquake that large. In other words, one magnitude-8 earthquake should occur every 1000 years or so. It is not surprising that we have not caught one of these large earthquakes on our instruments, since we have only been using them for 1/10 of that time interval. However, a city that has stood near that fault for 3000 years should have experienced three such large events during that time, along with many smaller events.

As the period for which we have been able to measure earthquake epicenters and magnitudes gets longer (it now stretches more than a century), we are beginning to see larger patterns emerge in the timing of large earthquakes, at least in some regions of the world. For instance, only 3 months after the enormous Sumatra earthquake of 26 December 2004, a very large earthquake, with a magnitude estimated at 8.7 as of this writing, occurred south of the first earthquake, along a section of the plate boundary that did not slip in the first event. Ironically, some of the first detailed scientific analyses of the Great Sumatran earthquake were just appearing in journals when the second earthquake occurred, and one article, published by Seth Stein and Emile Okal in the British journal *Nature*, warned of the continued danger on this southern section of the boundary (Stein and Okal, 2005). This is because we now understand that some earthquakes can actually increase the stress on sections of the same fault, or nearby faults, that did not slip initially (King et al., 1994). Thus, one earthquake can increase the chances of another earthquake nearby.

In fact, on a large fault, accumulated stress can be released in a series of large earthquakes, each one triggering another on an adjacent section of the fault a few months or years apart, so that the fault 'unzips' in steps until the stress has been released along its entire length. The whole process of gradual accumulation and subsequent release of stress then repeats, leading eventually to another series of earthquakes after centuries of quiescence. As a result, whole regions can be at significant earthquake risk, even in areas where modern seismicity is not significant.

Geophysicists have not reached a consensus on what to call this phenomenon, describing modern examples as 'earthquake sequences' (e.g., Ambraseys, 1970), 'earthquake migrations' (Mogi, 1968; Roth, 1988), 'progressive failures' (Stein et al., 1997), or, in places where many intersecting faults are involved, 'earthquake storms' (Nur and Cline, 2000).

An example of such an 'earthquake sequence' took place on the North Anatolian Fault during the twentieth century (e.g., Allen, 1975; Ambraseys, 1970; Stein et al., 1997; Wood, 1996). This 30-year 'sequence,' which is well known and studied among geophysicists, consisted of a series of earthquakes, all with magnitudes >5.6, in 1939, 1942, 1943, 1944, 1951, 1957, and 1967. The earthquakes progressed westward along the 1000 km long fault zone, releasing strain that had accumulated over the previous 200 or more years (**Figure 1**); this 'sequence' might now be extended to 60 years in length, in light of the August and November 1999 earthquakes that continued the earlier rupture pattern. The region affected by this sequence is home to many significant archaeological sites, including many cities of the Hittite empire, as well as the ruins of ancient Troy.

4.22.1.2 Limited Usefulness of History

While the precise timing, locations, and magnitudes of earthquakes cannot be predicted, and may follow complicated patterns, we can, given a sufficient history of instrumentally recorded earthquakes, estimate the mean recurrence interval for earthquakes of a given size on a particular fault. In practice, though, this technique has serious limitations, the greatest of which is the short time over which earthquakes have been instrumentally recorded. Seismometers have been in widespread use only since the early 1900s, and in many remote regions, the record is even more limited than that. Since the frequency–magnitude relation implies that the most devastating earthquakes will also be the least frequent, many regions of significant seismic hazards have not seen a major earthquake since the birth of modern seismology.

Thus, seismic hazard must often be based on crude historical data that do not include seismograms. Written records of historical earthquakes often include lists of cities where shaking was felt or where significant damage had to be repaired. By piecing together different accounts of such damage, we can crudely estimate the size and location of past earthquakes (Guidoboni et al., 1994). But written records are also a relatively recent thing, and the reliability of the oldest of these is suspect. Lists of affected areas are biased toward cities of political or religious importance and may entirely omit areas that were less populated or more isolated. Written records are also unlikely to survive for smaller earthquakes, which can be very destructive locally.

Another problem with historical earthquake records is that it can be difficult to determine which fault was responsible for an ancient earthquake. Estimates of damage were very subjective and were also influenced by politics and population density. Some earthquakes may well have happened on faults that have not yet been mapped and may represent completely unanticipated seismic hazard zones.

How then, besides waiting for more earthquakes to occur, can we extend our observation window to gain a more complete understanding of seismic risk? One of the most

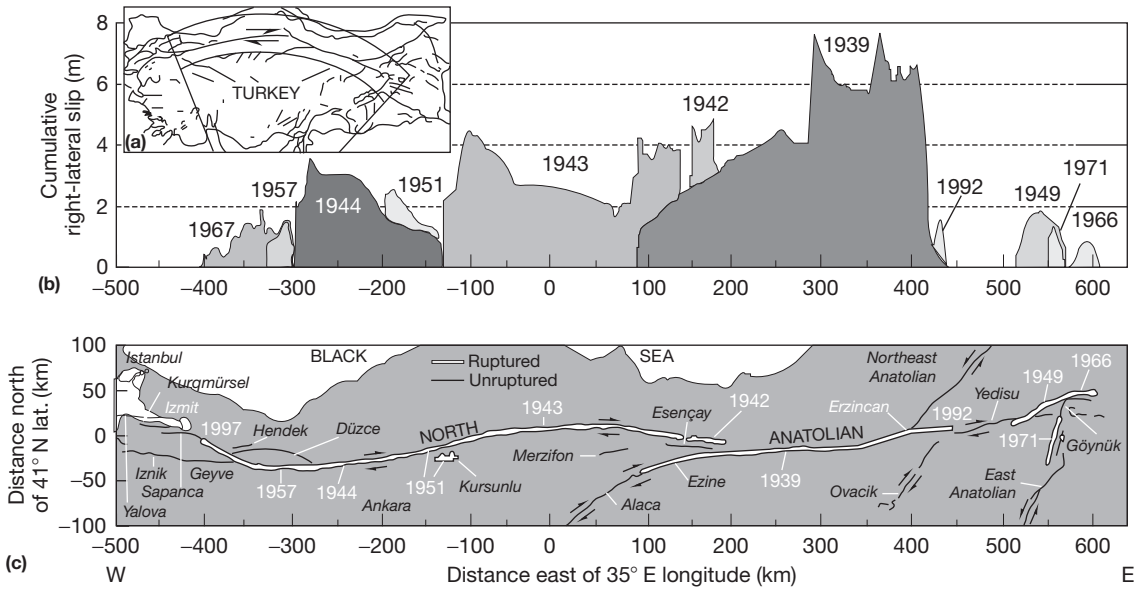


Figure 1 This map shows the locations of earthquakes that occurred in a sequence from 1937 to 1967 on the North Anatolian Fault. Reprinted from Stein R, Barka A, and Dietrich JH (1997) Progressive failure on the North Anatolian fault since 1939 by earthquake stress triggering. *Geophysics Journal International* 128(3): 594–604.

promising avenues for collaborative research is the relatively new field of archaeoseismology, the study of ancient earthquake damage using archaeological methods. This chapter will discuss the current state of this fledgling field and explore some of the obstacles scientists face in achieving a true synthesis of these two very different disciplines.

4.22.1.3 Effects of Tectonics on Sites of Human Occupation, Resonant Effects, and Vulnerability of Tells

Many sites of ancient human habitation are located in areas of significant seismic risk. The Middle East, Turkey, Greece, and Rome are renowned for their early cultures and often-magnificent ruins, and all are areas where earthquakes are common. The Americas, particularly south of the United States, had their own early urban cultures, many of which were in seismically active areas.

In fact, some ancient cities owe their locations in part to the effects of earthquakes. For example, the ancient walls of Mycenae in Greece were built upon a fault scarp (Maroukian et al., 1996), which afforded a steep, slick, natural foundation for the city's defenses (**Figure 2**). The builders of Mycenae were almost certainly unaware that this landform indicated a significant earthquake hazard, since the idea that earthquakes occur on faults is a recent one.

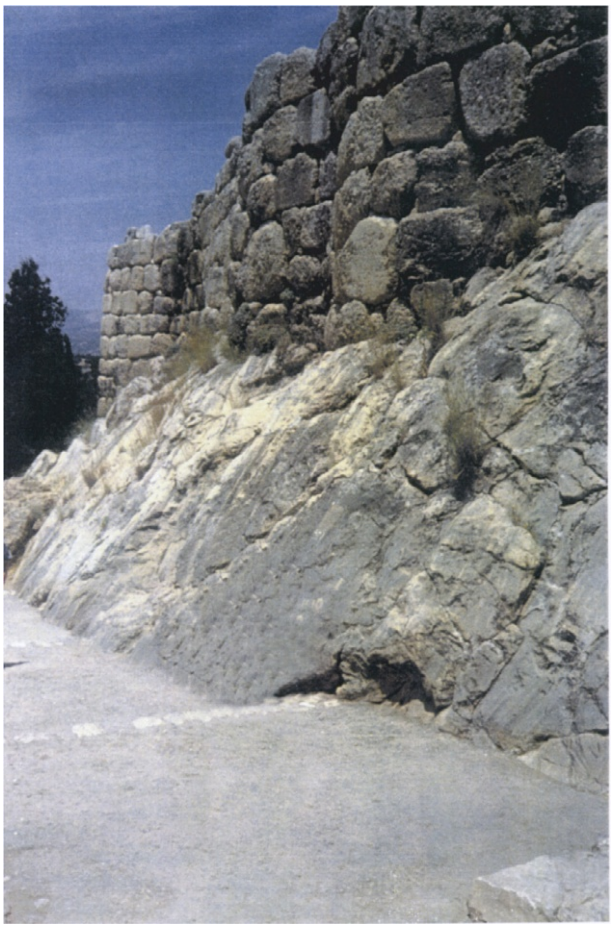


Figure 2 The defensive walls of ancient Mycenae were built atop a fault scarp, visible in this picture as the steeply sloped, light gray base of the wall.

The locations of other ancient cities are more subtly related to earthquakes, because the strategic value of earthquake-generated topography is a major factor that determines routes of trade and migration; cities that spring up along those routes may be subject to hazards from the same earthquakes that are responsible for their strategic importance. An example of this is Megiddo, in Israel (Aharoni, 1993; Kempinski, 1993). This ancient city has been a strategic battleground for millennia and, as recently as 1918, during the British mandate years. Accumulated motion on the Carmel–Gilboa fault system in this area resulted in a natural bottleneck in trade routes between Syria and the Mediterranean; Megiddo guards the major mountain pass through the Carmel mountains and thus has maintained political importance from the fifteenth century BC to the twentieth century AD. Multiple layers of destruction and rebuilding at this one site have left an artificial hill, called a tell by archaeologists. Some of the destruction events preserved in this tell were certainly the results of military action. However, the tectonic environment of the site makes it almost inescapable that some of the destruction is from ancient earthquakes (Nur and Ron, 1997).

Tells are diagnostic of important ancient sites in the Mediterranean region, and although the names and significance of many ancient cities have been lost to time, their mounds still stand, waiting to be excavated, preserving archaeological and geophysical secrets for future scientists to discover. The very nature of these tells makes them interesting to geophysicists for other reasons. The structure of the tells, with successive layers of construction built upon the unconsolidated rubble of previous layers, makes them particularly susceptible to earthquake damage, because of the weakness of unconsolidated fill when shaken. This is one of the effects that caused damage in the 1989 Loma Prieta, California, earthquake to be much greater in the San Francisco Marina district than in regions much closer to the epicenter. The damage in the Marina District was exacerbated by a high water table, resulting in saturated soil that could easily liquefy. This is less common in the arid Mediterranean region but can occur in coastal or riverside areas or in areas with natural springs.

Another phenomenon that tells exhibit during earthquakes is topography-related resonance when exposed to seismic waves, an effect that has been noted in modern earthquakes (Spudich et al., 1996). For example, during the Northridge, California, earthquake of 1994, 21 three-component geophones recorded the pattern of ground motion atop an elongated hill in Tarzana, California (**Figure 3**). Ground motion was generally much higher at the top of the hill than at the bottom of the hill and in the surrounding countryside. This pattern suggests that certain wavelengths of seismic energy can be focused and concentrated by the shape of the topography, creating a resonant effect that greatly increases the seismic hazard at the top of the hill. Thus, the resonant effects of the artificial topography, combined with the earthquake susceptibility of loose rubble foundations, may cause considerably more damage in tells than in the unaltered surrounding countryside, particularly when the tells are noncircular, as is true in most cases. This makes archaeoseismology potentially even more useful than paleoseismology in natural terrain.

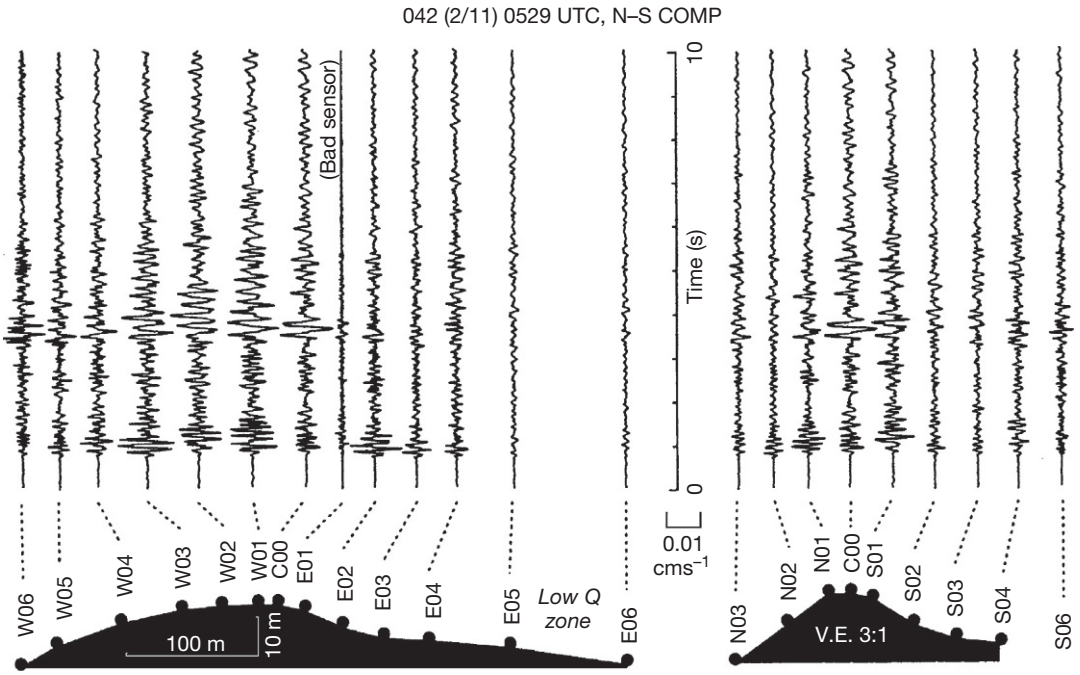


Figure 3 This is a plot of ground motion recorded on Tarzana hill by three-component geophones. Note that the shaking is much more severe atop the hill than on lower ground. This indicates a topographic resonance effect that is likely to contribute to increased damage atop tells. Reprinted from Spudich P, Hellweg M, and Lee WHK (1996) Directional topographic site response at Tarzana observed in aftershocks of the 1994 Northridge, California, earthquake: Implications for mainshock motions. *Bulletin of Seismological Society of America* 86(1B): S193–S208.

4.22.1.4 Communication Issues Between Disciplines

The differences between the damage levels in tells and in unaltered terrain environments are particularly important, because many currently inhabited cities in the Mediterranean region are built the same way the tells were – upon the rubble of past civilizations. Thus, modern cities in ancient regions share this extreme vulnerability to earthquakes. Unfortunately, many archaeologists who work in seismic hazard zones are completely unaware of these facts.

This is not surprising for, as far as I know, earthquake seismology is not a requirement of any archaeology degree program. Thus, the level of earthquake awareness among archaeologists is not much higher than that of the general educated public. Archaeology and geology share many characteristics, including the use of relative dating based on stratigraphy. However, archaeologists are not trained to recognize earthquake damage or even to know what sorts of damage can occur during an earthquake. Those who have never personally experienced an earthquake can have particularly unrealistic ideas about what earthquakes leave behind, expecting that they should find ground fissures or displaced faults cutting through the rubble if there has been any earthquake damage. When excavating, some archaeologists do not consider the possibility of earthquakes at all, and most of the rest consider earthquake damage to be an explanation of last resort when determining the cause of destruction in an ancient city.

Such reluctance is illogical, since we know that earthquakes remain an ever-present threat in any seismic hazard zone; in fact, if a city has existed in an earthquake-hazard zone for a significant period of prehistory, it is nearly inescapable that it should have experienced significant shaking. Earthquakes should be one of the primary suspects whenever a demolished city is found, unless there is specific evidence of human agents in the destruction. Even when evidence of human conflict is found amid the ruins, the possibility should be considered that enemies attacked opportunistically after an earthquake.

On the other hand, earthquake scientists are often ignorant of the conventions of archaeology. They may be unaware of the complex social and military interactions that have dominated a region, and since they see earthquakes as a scientific inevitability, they are certainly biased toward earthquake damage as an explanation for destruction. But perhaps a greater difficulty is that the jargon of seismology and the jargon of archaeology are different languages that often give different meanings to the same words. For instance, the word 'epicenter,' used in seismology, means the precise spot on the Earth's surface that is directly above the underground focus of an earthquake. However, the word is sometimes used loosely by archaeologists to mean the center of a region of destruction, with multiple epicenters sometimes being cited for the same earthquake. Archaeoseismologists need to learn some of the jargon of archaeology so that they can communicate the science behind earthquakes in a less ambiguous way. Then, it will be more possible to collaborate with archaeologists in uncovering traces of ancient earthquakes.

But what signs do earthquakes leave behind? The sections that follow examine some of the features of earthquake damage in archaeological remains in which archaeologists can document with traditional methods. They also explore special research methods that can complement traditional archaeology and that have had success in recent years resolving the prehistoric earthquake record in the Middle East.

4.22.2 Earthquake Traces in Archaeological Sites

It often can be very difficult to distinguish earthquake damage in archaeological sites from damage caused by slow decay over time or damage caused by human hands. This section will describe the various traces earthquakes leave behind. Documenting many of these types of damage from a single site during a single habitation level may implicate an earthquake as the most likely cause of destruction, even in the absence of historical data.

4.22.2.1 Effects of Earthquakes on Landforms

Perhaps the only unequivocal evidence for earthquakes is the presence of large-scale land deformation, associated with either faulting or wholesale rotation or displacement of the Earth's surface. There are several cases where such evidence has been uncovered in the archaeological record, but such cases are rare. However, this does not indicate, as some archaeologists have asserted, that earthquake damage is rare; most damage in earthquakes is associated with ground shaking, which may occur quite far from faults and may leave no evidence (beyond collapsed structures) after the shaking stops. Nevertheless, the dramatic nature of such land deformation makes it excellent evidence for earthquakes, when it is found.

4.22.2.1.1 Fault displacement

It is well known that earthquakes are the result of sudden displacement on faults in the Earth's crust. When earthquakes are large enough and shallow enough, that fault displacement may break the Earth's surface, resulting in ground displacement, which may cut through any structures that straddle the fault trace. The fault scarp described beneath the Mycenaean walls is unmistakable evidence that large earthquakes have occurred repeatedly in Mycenae. However, the fact that the wall was built atop the scarp only tells us that the earthquakes that formed the scarp predate the wall and that later earthquakes that activated the fault were insufficient to topple that wall (or the wall was subsequently repaired).

In tells, it is sometimes possible to bracket an earthquake between two successive layers of occupation. If a fault cuts and displaces the remains of one wall, but an undeformed wall has been built over the fault in the next layer, we know that the earthquake happened sometime between the building of the first wall and the building of the second.

An example of this is from the Crusader fortification of Ateret, by the banks of the Jordan River in northern Israel. The foundations of the fortress' walls have been displaced horizontally by motion on a fault that cuts through the walls; thus, the displacement must have occurred after the walls' construction in 1178 (and destruction by Saladin a year later). An early modern Muslim structure built within the ruins is also displaced, but whereas the Crusader walls are displaced by 2.1 m across the fault, the walls of the Muslim structure (dated somewhere during the Ottoman Turkish period of 1517–1917) are displaced by only 0.5 m. Trenches dug across the zone of deformation indicate that the slip occurred during several earthquakes (Ellenblum et al., 1998). By combining trench data with current geodetic measurements to rule out aseismic slip, Ellenblum et al. (1998) were able to determine that this fault is currently locked and that the displacement that has accumulated in the ruins has been accommodated by at least two large earthquakes. Thus, their research has contributed information toward the seismic risk estimates in the region, a region that includes heavily populated areas of Jordan, Israel, and Syria.

4.22.2.1.2 Elevated or submerged coastal land surfaces

While the example in the preceding text was an exciting and valuable contribution, most ruins are not directly bisected by active fault traces. Horizontal displacement is only detectable when a continuous wall or other features can be traced across the fault surface. However, when a town or structure was built on an ancient shoreline, any vertical displacement of the structures can be measured, even if they are far from the fault trace itself. There are many cases where archaeology has posited that a port or coastal fort has been destroyed during an earthquake. For example, the town of Dioscurias on the shore of the Black Sea shows evidence that it was destroyed by earthquakes in the first century AD (Nikonov, 1996). The external walls of the city have been found slightly tilted, partially collapsed, and submerged in Sukhumi Bay. Archaeological evidence was used to argue against gradual sea-level rise being responsible for this apparent subsidence, and geologic stratigraphy of the bay-bottom sediments was used to discount the possibility of a large landslide that moved the city into the bay. This is an excellent example of the use of interdisciplinary collaboration to implicate a major earthquake as the cause of this destruction, and it is particularly important in this region. Because local seismicity has been relatively quiet in modern times, the seismic hazard in this region has been estimated to be low to moderate. However, if Nikonov is correct, a very large earthquake, with Mercalli intensities of IX to X, has occurred in this region and therefore is likely to occur again. Additional paleoseismic and archaeoseismic work in the area is necessary to confirm or discount this hypothesis, one that has important consequences for the region's current inhabitants.

Submerged and uplifted coastal settlements have been more thoroughly studied in Greece, where a prolific ancient building tradition and a known history of earthquakes have made an ideal workshop for archaeoseismology. There have been many studies of apparent changes in coastal elevation in this region (e.g., Papageorgiou and Stiros, 1996; Pirazzoli, 1996; Pirazzoli et al., 1992), and the employment of geostratigraphy has been particularly rigorous. By using fossil evidence, especially from fragile organisms unlikely to survive a significant period of wave erosion, researchers can rule out mechanisms of gradual deformation or sea-level change in favor of sudden uplift or subsidence by earthquakes.

4.22.2.1.3 Tilting of horizontal surfaces

Besides displacing surfaces horizontally or vertically, fault motion can act to rotate surfaces, with the result that previously horizontal layers in the ground acquire an increasing dip over time. An example of this is seen in a small pull-apart basin that is part of the Dead Sea fault system in Israel. Located in this basin is Ubeidiya, thought by many to be the oldest known hominid construction on Earth (Bar Yosef, 1993). The archaeological structure in this case is the simplest possible (a floor made of

pebbles), put there by the appropriately named 'Pebble Man,' thought to be a member of the species *Homo erectus*. Some 1.5 million years ago, this floor was made by pressing pebbles, parts of stone artifacts, and even bones of now-extinct mammals into what was then the muddy bank of the proto-Jordan River.

When it was first excavated, it was not immediately clear what the surface was, since it is now inclined 60° from the horizontal. Geologists and archaeologists collaborating eventually realized that the surface had been horizontal, or nearly so, when it was built. The present tilt of the floor is due to its rotation about a horizontal axis during the 1.5 My since it was created, rotation due to the growth of the pull-apart basin, which in turn was caused by motion on the Dead Sea Transform and associated faults. Since modern data suggest that large earthquakes recur on average every 400 years or so in this area, some 3000–4000 such events must have occurred, each tilting the floor on average by the miniscule amount of just two-hundredths of a degree.

Rotation can also tilt vertical structures out of true or change the azimuths of walls. However, these two cases would be harder to detect in archaeological settings, since in the first case, nonvertical walls tend to collapse, and in the second, the original alignment of a wall is generally unknown.

4.22.2.2 Architectural Damage

The features described in the previous section are extreme, exciting, and generally strong indicators of earthquake damage. However, such ground deformation occurs only in the largest earthquakes and usually only relatively close to the responsible fault. Most damage in earthquake occurs because of the transient shaking of the ground, which leaves permanent evidence only on the aboveground structures, not in the ground itself. Thus, architectural damage, without associated fault traces or ground deformation, is what we should expect to find most often in archaeological excavations of earthquake-damaged cities. The challenge is to differentiate between earthquake destruction and either slow decay over time or sudden destruction by human action. There are several types of architectural damage that are suggestive of earthquake damage. When several of these signs are present in a single excavation, without contraindications that would implicate human destruction, earthquake should be a primary candidate for the destructive agent.

The archaeological tradition of carefully noting the location of every artifact found goes a long way toward making archaeological data useful to seismologists. However, when archaeologists are aware of the possibility of earthquakes and know the importance of directional alignment, angle, and regional patterns of collapse in interpreting earthquake damage, there is less chance that crucial data will be lost as the excavated fragments are removed for cataloging or restored to their original positions.

4.22.2.2.1 Collapsed walls with preferential axes of destruction

When walls are destroyed by invading armies, they tend to be destroyed haphazardly. Earthquakes, particularly smaller ones, can also cause this type of chaotic wall collapse, if the walls crumble slowly during an extended period of shaking. However, strong earthquakes often topple walls with the first arrival of the damaging shear waves, in which case the ruins will show a preferential direction of collapse. We find such damage in many different sites, including some layers at Troy, Jericho, and Mycenae. The ruins of storerooms at Masada in Israel illustrate this effect particularly well.

At the end of the Jewish War against the Romans in 70 AD, the stronghold of Masada, which towers over the Dead Sea in the Judean Desert, became the last free Jewish community in Judea (Yadin, 1966). It was eventually conquered by the Roman army, but its buildings were not destroyed until much later. There is strong evidence that this destruction happened in an earthquake, perhaps one that destroyed other sites in the area in 363 AD. One of the things that made it possible for Masada's people to withstand the Roman siege for as long as they did was their elaborate system of storerooms (Figure 4). The storehouses were designed as a series of long, narrow rooms, all parallel to each other, with walls of unreinforced stone 11 ft high. This type of construction is particularly vulnerable to shaking perpendicular to the long direction of the storerooms, since the long walls have no bracing along most of their length. The shorter walls are braced at every intersection with the long walls, at least until the long walls collapse.

The ruins at Masada show that the storehouse walls apparently collapsed as a single unit in one direction, with each wall falling into the storeroom next to it. The rows of stones seen in the aerial photo in Figure 5, which was taken before excavation (and restoration of the walls) began, are the kind of marker that we have come to look for as evidence of an earthquake.

4.22.2.2.2 Aligned fallen columns

Not all building styles incorporate columns, but those that do are goldmines for archaeoseismology. Perched on their narrow

Figure 4 This view of the reconstructed storerooms of Masada shows the long, unsupported walls that are susceptible to earthquake damage.

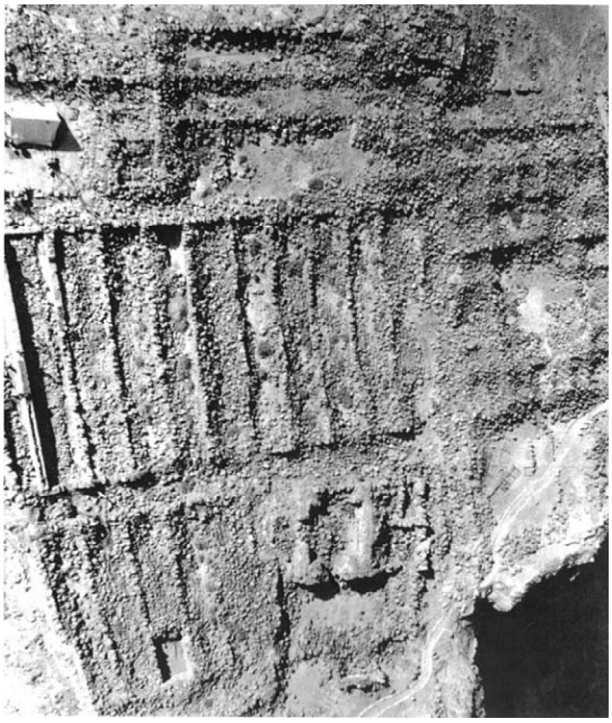


Figure 5 This picture, taken before the reconstruction of Masada, shows how the walls of the parallel storerooms collapsed into the adjacent rooms, leaving rows of bricks that are diagnostic of earthquake collapse. Reproduced from Masada: Herod's Fortress and Zealot's Last Stand by Yigael Yadin, © 1966 by Yigael Yadin, with permission of Random House, Inc.

bases, columns are very vulnerable to earthquake damage. In Greek designs, these columns were often supported on top only by the next course of stone, and in Roman construction, only by light wooden structures. Thus, when such columns fell in strong earthquakes, they often fell together in a very distinctive way, as in the ruins of Susita in northern Israel, shown in Figure 6. The most striking feature here is that all the columns are aligned in the same direction, indicating that they fell in response to single impulse. The impulse here was the 749 AD earthquake that hit the Sea of Galilee region. Since the bases of the columns are not tilted, which could indicate a slow tilting followed by collapse, this damage almost certainly happened during an earthquake. Had the columns been destroyed by human action, it would have been very difficult for the invaders to achieve such precise alignment. Examples of such systematic column collapse are widespread, including Knidos in southwestern Turkey; Jerash, Petra, and Pella in Jordan; Beit She'an in Israel; Chichén Itzá in Mexico; and Selinunte in Sicily, where the association with an earthquake is historically documented.

Another more subtle feature of Figure 6 is the fact that the bases of the columns do not line up with the direction of collapse. Instead, the lines that connect the bases to the columns that were perched on them are all parallel to one another in a slightly different direction. There are two possible explanations for this: either the columns were bucked sideways off their bases and then fell during subsequent shaking, or they fell and then rolled out of alignment with the bases during subsequent shaking.

Figure 6 These columns were discovered in the ruins of Susita, overlooking the Sea of Galilee. Note how the columns all fell in approximately the same direction, yet they are no longer aligned with their bases. These are signs that the columns collapsed in an earthquake, probably one that occurred in 749 AD.

4.22.2.2.3 Slipped keystones

Sometimes, an earthquake does not topple a wall, but only partly damages it. In this case, the remains of the wall preserve telltale signs of earthquake action. Arches are particularly good places to look for this evidence.

Arches are used to allow large openings in stone walls while still supporting the weight of stone above. The key to the strength of this structure, which could be built without mortar or pins, was allowing the weight of the stones to compress the arch, pushing each stone firmly against its neighbor. Unless a stone completely disintegrated under the pressure, a well-designed arch would stand indefinitely under its own weight. However, if a single of the wedge-shaped stones, or 'voussoirs,' were removed from the perimeter of the arch, the entire structure could collapse.

In ordinary circumstances, this poses no problem. The immense weight of a masonry structure holds the stones firmly in place. However, in the extraordinary circumstance of an earthquake, the Earth's motion wields the structure's own weight against it, effectively throwing the whole edifice around. Moderate and large earthquakes can result in very high ground acceleration, occasionally even exceeding the acceleration of gravity, particularly in areas where the local topography or geology amplifies ground shaking (as in tells). In other words, an earthquake can effectively cancel gravity, which would obviously have disastrous effects for gravity-dependent structures. However, even when the motion is less severe, structures are subjected to both strong horizontal forces and vertical forces acting against gravity.

The key to the weakness of arch-based architecture is the ease with which the base of the arch can slide horizontally, especially if partially deprived of gravitational compression. Typically, even in the second half of the twentieth century, the bases of arches were not pinned in any way to the structures beneath them. This means that, under the enormous horizontal forces generated in an earthquake, the base of the arch can simply slip sideways on its foundation, relieving the compression that holds the arch together and causing total collapse. Because there was no concept of seismic hazard in ancient

times, that is, no idea that earthquakes could and should be expected to repeat, and no way to grasp the huge forces involved in earthquakes, no thought was given to pinning the structures against horizontal motion as an antiseismic precaution. In fact, when exactly such earthquake damage occurred at Stanford University during the 1906 San Francisco earthquake, the displacement was simply corrected without pinning the arches' bases; as a result, the same arches were damaged again in the 1989 Loma Prieta earthquake.

Most arches that were hit by sufficiently large past earthquakes simply collapsed, and all we find in the excavated rubble is a heap of stones that made up the arch. Unless the whole wall toppled neatly outward, like in Figure 7, the chaotic remains of collapsed arches are not very useful indicators of earthquake damage. They are too difficult to distinguish from collapse due to slower Earth movements or destruction by violence. However, earthquakes sometimes stop short of collapsing the building, and one or more of the voussoirs in an arch drop down until it is again pressed tight against its neighbors.

Figure 8 shows an example of this from Qala'at Namrud, an Arab and Crusader fortification on the main road from Galilee to Damascus. On the left side, indicated by the arrows in the photograph, the whole base of the doorway slipped sideways, loosening the stones in the arch. Subsequently, several of the voussoirs in the arch dropped down, and the arch was again under compression, leaving the doorway a bit wider and shorter than the original architect had intended.

Another example can still be seen in Rome, in the Colosseum, which was partly destroyed by an earthquake in 1349. Figure 9 shows some of the arches that remain. The arches that would once have stood to the left of these collapsed in the destructive earthquake, but the leftmost one in this photo did not. The movement was arrested just in time, as we can see in the close-up. The voussoirs slipped only partially, and the arch survived.

4.22.2.3 Skeletons of Trapped Humans and Domestic Animals

One of the challenges in distinguishing earthquake damage from other sorts of damage is the difficulty of showing that the collapse was sudden and unexpected. It is a sure sign of this when skeletons of humans or domestic animals are found trapped beneath collapsed walls, especially when they are found in defensive postures. Trapped skeletons have been found in archaeological sites around the world, including Mycenae, Tiryns, Midea, Thebes, Crete, and the Menelaion in

Figure 7 These arches are from Beit She'an, 45 km south of Susita, shown in Figure 6. These arches collapsed out into the street, probably in the same earthquake as destroyed Susita in 749 AD.

Figure 8 This is an example of a slipped voussoir in the Crusader fortification of Qala'at Namrud. Note how the base of one side of the arch slipped outward, allowing the voussoir to drop.

Figure 9 These arches are from the outer wall of the Colosseum in Rome. Once, there were more arches to the left of the leftmost one in this photo, but they collapsed in an earthquake. Note the chaotic state of the keystones in the leftmost remaining arch.

Greece (Åström and Demakopoulou, 1996; Catling, 1981; Kilian, 1996; Mylonas, 1966; Mylonas-Shear, 1969, 1987; Sakellarakis and Sapouna-Sakellaraki, 1981; Sampson, 1996) and Dor, Petra, Jerash, and Pella in the Middle East (Amiran et al., 1994; Browning, 1982; Stewart, 1993; Stucky, 1990).

The oldest known examples of skeletal remains of earthquake victims come from Shanidar Cave in Iraq. In that cave, there are multiple layers in which Neanderthal skeletons are crushed under boulders that fell from the ceiling of the cave, presumably in earthquakes. Interestingly, the excavator of the cave, Ralph Solecki, experienced an earthquake inside the cave while excavating (Solecki, 1971).

The ancient port of Kourion, located on the southern coast of Cyprus, has revealed compelling evidence of earthquake victims. In 1934, archaeologist J. F. Daniel unearthed two crushed human skeletons dubbed 'Romeo and Juliet' beneath fallen rubble of their house, which also held coins dated up to 365 AD and crushed wares (Soren, 1988a). The early 1980s excavation by David Soren found several human skeletons crushed by falling walls and roofs of separate buildings. Some skeletons were curled into fetal postures. A room revealed a man in his thirties, crushed under rubble with his hands spread to cover his head. Another room, apparently a stable, held the bones of a mule still tethered to a stone trough. Intertwined with the animal's skeleton is that of a 13-year-old girl, caught perhaps trying to calm her frightened beast. One remarkable find is of a young family of three (**Figure 10**), trapped under the ruins of their bedroom (Soren, 1988a,b).

Figure 10 Kourion, Cyprus: A young couple and their 18-month-old child crushed under blocks from the ceiling of their bedroom, some weighing as much as 300 pounds. Reprinted from Soren D (1988) The day the world ended at Kourion: Reconstructing an ancient earthquake. *National Geographic* 174(1): 30–53, with permission. Credit: Martha Cooper/National Geographic Image Collection.

These archaeological findings are associated with the historical earthquake of 365 AD.

4.22.2.4 Fire

Fire is one of the types of damage that archaeologists generally attribute to invaders. This is because archaeologists and the general public have a misconception about the firestorms that often follow severe earthquakes. In modern times, such fires are usually caused by ruptured natural gas lines and downed electrical wires, so the popular view is that such fires would not have occurred in antiquity, before the advent of such modern hazards. However, in antiquity, heating, cooking, and lighting were all accomplished by open flame, which in itself is a significant hazard. In an earthquake, fires burning in a hearth or lamp can be upset or scattered into a dwelling, where combustible materials can ignite.

Another argument against fires following ancient earthquakes is that the building materials used in ancient buildings (stone, mud brick, and tile) were not combustible. This argument is illogical; if the buildings were so fireproof that an earthquake could not ignite them, then how did the enemies manage to do so? Did enemies haul in all the fuel necessary to burn an entire settlement?

While it is true that the remains we find in archaeological sites are generally incombustible, it is not true that combustible materials were not used in ancient buildings; it is just that combustible materials do not generally survive millennia of burial. Roof support timbers were often used in all but the grandest of stone structures, and furnishings were generally made of wood and cloth. And the fact is, there are plenty of known cases of fires following earthquakes that predate modern gas and electrical utilities. The fire that destroyed Lisbon, Portugal, in 1755 (British Historical Society of Lisbon, 1990) was one of the most dramatic examples; it burned for 6 days after the earthquake and was responsible for the loss of uncounted treasures of art and history in that city.

Therefore, evidence of widespread fire in an archaeological site is just as likely to be due to an earthquake as to human action and certainly should not be used as an argument against earthquake destruction.

4.22.3 Techniques in Archaeoseismology

The field of archaeoseismology is a relatively new one, and it is by nature highly collaborative. Despite the difficulties in cross-disciplinary communication, scientists from the fields of geology, geophysics, archaeology, paleontology, and even forensic anthropology have made great strides toward integrating their varied perspectives into a coherent approach to ancient earthquakes. There are several techniques that have been employed to that end.

4.22.3.1 Informing Archaeological Methods with a Seismology Perspective

The first and most important of archaeoseismological techniques is simply one of disseminating knowledge and changing perspectives. The standard techniques of archaeology are ideally

suited to reconstructing an earthquake scenario, as long as the archaeologist is aware of the types of data that seismologists need. Not only locations of artifacts but also their orientations and the structure of the soil in which they are buried are important. These are data that archaeologists can record, and often do, in the course of a normal excavation. However, there are instances where specialized expertise is required, which may warrant the involvement of an outside expert.

4.22.3.2 Intersection of Archaeology and Paleoseismology

One case where additional expertise may be necessary is when there is evidence of a fault structure cutting through an archaeological site. In this case, the involvement of paleoseismology can make an excavation much more informative.

Paleoseismologists specialize in recognizing traces of ancient earthquakes in subsurface deposits, using trenching methods to examine layers and how they have been displaced relative to each other. They are familiar with how earthquakes disturb sediments and may be able to pinpoint relative dates for earthquakes more accurately than an archaeologist who is not familiar with examining earthquake traces. Paleoseismologists are also more likely to make use of microscopic biological or paleontological data than are archaeologists, who are more likely to use pottery and other cultural indicators, and thus may be able to produce an independent relative date for comparison with the archaeological date.

Another paleoseismological method that has shown promise for informing archaeological digs is the analysis of sea- or lake-bottom sediments for evidence and dating of earthquakes. When an earthquake shakes the water-saturated, unconsolidated sediments on a sea or lake floor, it creates layers of chaotic, mixed sediment called seismites. In an analysis of Dead Sea seismites (Ken-Tor et al., 2001), radiocarbon dating of organic matter in the disturbed layers allowed seismites to be correlated to eight historical earthquakes, dating from 64 BC to 1927 AD. In the Dead Sea region, where the historical record is very long, such correlation can help scientists estimate the size of these ancient earthquakes and determine what archaeological sites might have been affected around the Dead Sea; it can give clues to possible earthquake epicenters as well. In a region without such long records, seismite information might reveal entirely unknown seismic hazards.

4.22.3.3 Forensic Anthropology

Another case where outside expertise is indicated is in determining the cause of death when a skeleton is found in the ruins. In the case of skeletons found under collapsed walls, it may be crucial to determine whether the victim was trapped by falling debris in an earthquake or died by other means and was left with its skeleton eventually being buried in the slow decay of the building. The field of forensic anthropology is uniquely equipped to deal with these sorts of questions, using medical knowledge of how fresh bone fractures differently than dry bone and how the presence of soft tissues around and inside bone can change how it responds to sudden force, burial, or fire.

The discovery of a female skeleton buried beneath a wall in the coastal Israeli town of Dor prompted archaeologist Andrew Stewart (Stewart, 1993) to consult an expert in bone fracture analysis to determine whether the falling wall could have killed her or whether she might have been already dead when the wall fell. While the distinction cannot be made with that precision, Dr. Patricia Smith at Hadassah Medical Center in Jerusalem found evidence of a particular kind of breakage, a spiral fracture, that only occurs in fresh bone. This made it clear that she was either alive or fairly freshly dead when the falling wall broke her bones. Further evidence can be obtained by reconstructing the victim's position to determine whether the fractures occurred *in situ* or while the victim was standing erect. These are generally determinations that an archaeologist or geologist is not qualified to make.

Another instance where forensic knowledge was able to extract more data from a site than archaeology or geology could do alone is the excavation of a Minoan temple from about 1700 BC in Archanes, Crete. Excavators of this site found several skeletons within the collapsed building, but in this case, they uncovered a particularly shocking scene. As the main room of the temple was excavated, archaeologists discovered a priest beside an altar, with his sacrificial victim still bound before him. It had been thought that altars from this period were used for the sacrifice of bulls, not of young men as was the case in this scene.

From the archaeological evidence, either during the ceremony or immediately before or after the sacrificial act, an earthquake struck, collapsing the building. The earthquake was followed by an intense fire, which charred the entire site, including the bones of the people trapped within. Forensic study of the bones on the altar was able to answer the timing question that archaeology alone could not resolve.

According to the forensic anthropologist, bones that are burned in the presence of blood turn black, while if the corpse is drained of blood, the bones would remain white. In the sacrificial victim, who was lying on his side, the bones in the uppermost side of his body were white, while the lower bones were black. This indicates that much of his blood had been drained before his body burned (Sakellarakis and Sapouna-Sakellaraki, 1981).

4.22.3.4 Numerical Simulation

One valuable piece of information that scientists hope to gain from archaeoseismology is a magnitude estimate for the earthquake that destroyed the site in question. There are many factors that determine the damage caused by an earthquake, including the distance from the epicenter, the depth of the earthquake's focus, the duration of shaking, the structure of the subsurface, and the type of architecture. All of these factors combine to create a seismic intensity, which is simply a numerical description of the damage level.

Given an archaeological site with, say, huge collapsed columns, determining the amount of shaking required to topple the columns is not an archaeological or seismological problem, but rather an engineering problem. Numerically simulating the effects of shaking on engineering structures is a complex problem, but it can yield some information about possible earthquake magnitudes (e.g., Papastamatiou and Psycharis, 1996).

Other simulations can reveal how the structure of subsurface sediments can intensify shaking during an earthquake. This technique was used to show why the damage to Rome's

Colosseum in the 1349 earthquake was so much more severe on its southern edge than in the rest of the structure. The entire southern half of the exterior collapsed, while the rest remains standing to this day (Figure 11). Studies of the ground beneath the Colosseum show that the southern portion of the structure is built on a buried river channel, which is filled with unconsolidated, water-saturated sediments, while the rest of the structure is underlain by more stable layers. A numerical simulation of the response of the structure beneath the stadium confirmed that resonance of earthquake waves within the buried river channel would greatly amplify the shaking of the ground above (Moczo et al., 1995). In terms of seismic hazard estimates, this discovery highlights the importance of using architectural remnants of ancient buildings to inform scientists of possible hidden seismic hazards in the subsurface. Besides the damage pattern in the Colosseum, there is no visible evidence at ground level of this buried river channel; thus, the pattern of ancient earthquake damage alerted modern seismologists of a seismically vulnerable structure that also threatens more modern buildings in the same area.

4.22.4 Mutual Benefits of Archaeoseismology

4.22.4.1 Benefits to Seismology

Seismology can gain a great deal by incorporating data from archaeology (Nur, 2002). The most notable benefit is the ability to extend the observation window for seismic events to include prehistoric events. This can of course be done with paleoseismology as well, but the ongoing excavations of so many archaeological sites around the world are a potential wellspring of seismological information that now largely goes untapped.

In addition to extending our observation window in time, archaeology can also give us better spatial coverage, since many places for which we have few historical records have archaeological sites that could yield seismic information. This is particularly true in the Americas, where our historical record is particularly poor prior to European contact.

Finally, as shown by the numerical studies of the Colosseum, archaeology can fine-tune our seismic hazard maps, revealing seismically weak ground structure that may be invisible from the surface. Areas that suffered extreme damage in ancient times can indicate places where we should carry out subsurface seismic studies, to determine whether resonance or other complicating effects make the seismic hazard higher than in the surrounding region.

Figure 11 An aerial view of the Colosseum in Rome.

4.22.4.2 Benefits to Archaeology

The obvious benefit of archaeoseismology to archaeology is a clearer understanding both of what earthquake destruction looks like after the fact and of where earthquake damage is most likely. There are certainly many instances where earthquake damage has been ascribed to human action, not because there is evidence of warfare or intentional destruction, but because that is the usual assumption. By being aware of earthquake hazards and being trained to recognize earthquake damage, archaeologists can incorporate seismological causes into their hypotheses without the feeling that they are resorting to a *deus ex machina* in the absence of any other ideas.

Another potential benefit of archaeoseismology stems from earthquakes' randomness and independence from human factors. Seismic waves travel much faster than human migrations or cultural diffusion. Thus, if the region affected by a particular ancient earthquake can be established with reasonable certainty, that earthquake could be used as a correlation layer. This would allow relative dating with unprecedented precision among all the sites in the earthquake-affected region, regardless of the dominant cultures of the sites and regardless of whether the sites had any commerce with one another.

Finally, although ancient earthquakes were a tragedy to those they affected, they are a boon for archaeologists. The excavation of tombs teaches us much about how ancient peoples died; the excavation of abandoned cities teaches us much about how they built; the excavation of ancient battlegrounds teaches us much about how they fought. But only sudden, unexpected catastrophes like earthquakes can preserve snapshots of how they actually lived their day-to-day lives.

References

Aharoni Y (1993) Megiddo. In: Stern E (ed.). *New Encyclopedia of Archaeological Excavations in the Holy Land*, vol. 3, pp. 1003–1012. New York: Simon and Schuster.

Allen CR (1975) Geological criteria for evaluating seismicity. *Geological Society of America Bulletin* 86: 1041–1057.

Ambraseys NN (1970) Some characteristic features of the Anatolian fault zone. *Tectonophysics* 9: 143–165.

Amiran DHK, Arieh E, and Turcotte T (1994) Earthquakes in Israel and adjacent area: Macroseismic observations since 100 BCE. *Israel Exploration Journal* 44: 260–305.

Åström P and Demakopoulou K (1996) Signs of an earthquake at Midea? In: Stiros S and Jones RE (eds.) *Archaeoseismology*, pp. 37–40. Athens: The British School at Athens, Fitch Laboratory Occasional Paper 7.

Bar Yosef O (1993) Ubeidiya. In: Stern E (ed.) *The New Encyclopedia of Archaeological Excavations in the Holy Land*, pp. 1487–1488. Carta, Jerusalem: The Israel Exploration Society.

British Historical Society of Lisbon (1990) Lisbon recalled: All Saints' Day, 1 November 1755. *Terramoto de 1755. Testemunhos Britanicos*, British account Lisboa, pp. 670–672.

Browning I (1982) *Jerash and the Decapolis*. London: Chatto & Windus Ltd.

Catling HW (1981) *Archaeology in Greece, 1980–81*. Archaeological Reports for 1980–81, 3–48, London: The Society for the Promotion of Hellenic Studies.

Ellenblum R, Marco S, Agnon A, Rockwell T, and Boas A (1998) Crusader castle torn apart by earthquake at dawn, 20 May 1202. *Geology* 26(4): 303–306.

Guidoboni E, Comastri A, and Traina G (1994) *Catalogue of Ancient Earthquakes in the Mediterranean Area up to the 10th Century*. Phillips B (trans.), Rome: Insituto Nazionale di Giofisica.

Kempinski A (1993) *Megiddo: A City-State and Royal Centre in North Israel*. Munich, Germany: Verlag C. H. Beck.

Ken-Tor R, Agnon A, Enzel Y, Stein M, Marco S, and Negendank JFW (2001) High-resolution geological record of historic earthquakes in the Dead Sea basin. *Journal of Geophysical Research* 106(B2): 2221–2234.

Kilian K (1996) Earthquakes and archaeological context at 13th century BC Tiryns. In: Stiros S and Jones RE (eds.) *Archaeoseismology*, pp. 63–68. Athens: The British School at Athens, Fitch Laboratory Occasional Paper 7.

King GCP, Stein RS, and Lin J (1994) Static stress changes and the triggering of earthquakes. *Bulletin of Seismological Society of America* 84: 935–953.

Maroukian H, Gaki-Papanastassiou K, and Papanastassiou D (1996) Geomorphologic-seismotectonic observations in relation to the catastrophes at Mycenae. In: Stiros S and Jones RE (eds.) *Archaeoseismology*, pp. 189–194. Athens: The British School at Athens, Fitch Laboratory Occasional Paper 7.

Moczo P, Rovelli A, Labak P, and Malagnini L (1995) Seismic response of the geologic structure underlying the Roman Colosseum and a 2-D resonance of a sediment valley. *Annali di Geofisica* 38(5–6): 939–956.

Mogi K (1968) Migration of seismic activity. *Bulletin of the Earthquake Research Institute* 46: 53–74.

Mylonas G (1966) *Mycenae and the Mycenaean Age*. Princeton: Princeton University Press.

Mylonas-Shear I (1969) *Mycenaean Domestic Architecture*. Ph.D. Thesis, Ann Arbor, University of Michigan.

Mylonas-Shear I (1987) *The Panagia Houses at Mycenae*. Philadelphia: University Museum Press.

Nikonov A (1996) The Disappearance of the ancient towns of Dioscuria and Sebastopolis in Colchis on the Black Sea: A problem in engineering geology and palaeoseismology. In: Stiros S and Jones RE (eds.) *Archaeoseismology*, pp. 195–204. Athens: The British School at Athens, Fitch Laboratory Occasional Paper 7.

Nur A (2002) Earthquakes and archaeology. In: Lee WHK, Kanamori H, Jennings PC, and Kisslinger C (eds.). *International Handbook of Earthquake and Engineering Seismology*, 81A, pp. 765–774. London: Academic Press.

Nur A and Cline EH (2000) Poseidon's horses: Plate tectonics and earthquake storms in the Late Bronze Age Aegean and Eastern Mediterranean. *Journal of Archaeological Science* 27(1): 43–63.

Nur A and Ron H (1997) Armageddon's earthquakes. *International Geology Review* 39: 532–541.

Papageorgiou S and Stiros SC (1996) The harbour of Aigeira (North Peloponnese, Greece): An uplifted ancient harbour. In: Stiros S and Jones RE (eds.) *Archaeoseismology*, pp. 211–214. Athens: The British School at Athens, Fitch Laboratory Occasional Paper 7.

Papastamatiou D and Psycharis I (1996) Numerical simulation of the seismic response of megalithic monuments: Preliminary investigations related to the Apollo temple at Vassai. In: Stiros S and Jones RE (eds.) *Archaeoseismology*, pp. 225–236. Athens: The British School at Athens, Fitch Laboratory Occasional Paper 7.

Pirazzoli PA (1996) Uplift of ancient Greek coastal sites: Study, methods and results. In: Stiros S and Jones RE (eds.) *Archaeoseismology*, pp. 237–244. Athens: The British School at Athens, Fitch Laboratory Occasional Paper 7.

Pirazzoli PA, Ausseil-Badie J, Giresse P, Hadjidaki E, and Arnold M (1992) Historical environmental changes at Phalasarna harbor, West Crete. *Geoarchaeology* 7: 371–392.

Roth F (1988) Modeling of stress patterns along the western part of the North Anatolian fault zone. *Tectonophysics* 152: 215–226.

Sakellarakis Y and Sapouna-Sakellaraki E (1981) Drama of death in a Minoan temple. *National Geographic* 174: 205–222.

Sampson A (1996) Cases of earthquakes at Mycenaean and Pre-Mycenaean Thebes. In: Stiros S and Jones RE (eds.) *Archaeoseismology*, pp. 113–118. Athens: The British School at Athens, Fitch Laboratory Occasional Paper 7.

Solecki RS (1971) *Shanidar: The First Flower People*. New York: Knopf.

Soren D (1988a) *Kourion: The Search for a Lost Roman City*. New York: Anchor Books.

Soren D (1988b) The day the world ended at Kourion: Reconstructing an ancient earthquake. *National Geographic* 174(1): 30–53.

Spudich P, Hellweg M, and Lee WHK (1996) Directional topographic site response at Tarzana observed in aftershocks of the 1994 Northridge, California, earthquake: Implications for mainshock motions. *Bulletin of Seismological Society of America* 86(1B): S193–S208.

Stein R, Barka A, and Dietrich JH (1997) Progressive failure on the North Anatolian fault since 1939 by earthquake stress triggering. *Geophysics Journal International* 128(3): 594–604.

Stein S and Okal EA (2005) Speed and size of the Sumatra earthquake. *Nature* 434: 581–582.

Stewart A (1993) A death at Dor. *Biblical Archaeology Review* 19(30–36): 84.

Stucky RA (1990) Schweizer Ausgrabungen in ez Zantur, Petra: Vorbericht der Kampagne 1988. *Annual of the Department of Antiquities of Jordan* XXXIV: 249–283.

Wood M (1996) *In Search of the Trojan War*, 2nd edn. Berkeley: University of California Press.

Yadin Y (1966) *Masada: Herod's Fortress and the Zealot's Last Stand*. New York: Random House, Inc.

4.23 Earthquake Hazard Mitigation: New Directions and Opportunities

RM Allen, University of California Berkeley, Berkeley, CA, USA

© 2007 Elsevier B.V. All rights reserved.

This chapter is reproduced from the previous edition, Volume 4, pp. 607–647, © 2007 Elsevier B.V.

4.23.1	**Introduction**	593
4.23.2	**Recognizing and Quantifying the Problem**	594
4.23.2.1	Forecasting Earthquakes at Different Spatial and Temporal Scales	594
4.23.2.2	Global Seismic Hazard	595
4.23.2.3	Changing Seismic Risk	596
4.23.2.3.1	Earthquake fatalities since the 1900	596
4.23.2.3.2	Concentrations of risk	598
4.23.2.4	Local Hazard and Risk: The San Francisco Bay Area	599
4.23.2.4.1	The San Francisco Bay Area	599
4.23.2.4.2	Earthquake probabilities	599
4.23.2.4.3	Future losses	601
4.23.3	**The 'Holy Grail' of Seismology: Earthquake Prediction**	602
4.23.4	**Long-Term Mitigation: Earthquake-Resistant Buildings**	603
4.23.4.1	Earthquake-Resistant Design	603
4.23.4.1.1	Lateral forces	603
4.23.4.1.2	Strong-motion observations	603
4.23.4.1.3	Strong-motion simulations	604
4.23.4.1.4	New seismic-resistant designs	604
4.23.4.2	The Implementation Gap	607
4.23.4.2.1	The rich and the poor	607
4.23.4.2.2	The new and the old	607
4.23.5	**Short-Term Mitigation: Real-Time Earthquake Information**	608
4.23.5.1	Ground Shaking Maps: ShakeMap and Beyond	608
4.23.5.1.1	ShakeMap	609
4.23.5.1.2	Rapid finite-source modeling	609
4.23.5.1.3	Applications of ShakeMap	610
4.23.5.1.4	Global earthquake impact: PAGER	612
4.23.5.2	Warnings Before the Shaking	614
4.23.5.2.1	S-waves versus P-waves	614
4.23.5.2.2	Single-station and network-based warnings	615
4.23.5.2.3	Warning around the world	615
4.23.5.2.4	ElarmS in California	616
4.23.5.2.5	Warning times	619
4.23.5.2.6	Future development	620
4.23.5.2.7	Benefits and costs	621
4.23.6	**Conclusion**	621
Acknowledgments		622
References		622

4.23.1 Introduction

Few natural events can have catastrophic consequences of earthquakes, yet evidence abounds for repeating disasters in the same location. Archaeological studies point to the recurring destruction of Troy, Jericho, and Megiddo in the Mediterranean and the Middle East, and, in the New World, debris from the 1906 San Francisco earthquake was found beneath the destruction caused by the 1989 Loma Prieta earthquake in San Francisco's Marina District.

Historical examples illustrate the sociopolitical impact of earthquakes. In 464 BCE, a powerful earthquake beneath the ancient Greek city Sparta led to the rebellion of Spartan slaves. According to Aristotle's Politics (1269a37–b5), these slaves were "like an enemy constantly sitting in wait for the disasters of the Spartans." The devastation that Sparta suffered from the earthquake offered them the perfect opportunity. The rebellion, which lasted for 10 years, limited Sparta's ability to check the growth of Athenian power in Greece and also led to the dissolution of the Spartan–Athenian alliance formed some 30 years earlier in the face of the Persian threat.

More recently, another natural disaster destroyed much of New Orleans and the Gulf coast of Louisiana. Few believed a natural hazard could be so devastating to a modern wealthy city,

yet Hurricane Katrina flooded 80% of the city, much of which is below sea level, and destroyed over 300 000 housing units in August 2005. Despite a warning of the impending hurricane several days in advance, over 1800 people were killed. One year later, the population of the city is less than 50% of its previous level and it is clear that many will not return.

The challenge of natural hazard reduction generally, and earthquake hazard mitigation in particular, is the long return interval of these events. The infrequency of large seismic events provides only a limited data set for the study of earthquake impacts on modern cities, and the uncertainty as to when the next event will occur often places earthquake mitigation low on the priority list. The fields of seismology and earthquake engineering are also relatively juvenile, having only developed out of large destructive earthquakes at the end of the nineteenth and beginning of the twentieth centuries. Still, there has been considerable progress. Your chance of being killed in an earthquake is a factor of 3 less than it was in 1900.

Yet earthquakes account for 60% of natural hazard fatalities today (Shedlock and Tanner, 1999). The number of people killed in earthquakes continues to rise in poorer nations, and the cost of earthquakes continues to rise for rich nations. The global population distribution is changing rapidly as underdeveloped nations continue to grow most rapidly in cities that are preferentially located in seismically hazardous regions. There has not yet been a large earthquake directly beneath one of these megacities, but when such an event occurs, the number of fatalities could exceed 1 million (Bilham, 2004).

This chapter considers seismic hazard mitigation. First, we evaluate the hazard and risk around the globe to identify where mitigation is necessary. Next, we consider the topic of earthquake prediction that is often called upon by the public as the solution to earthquake hazard. Instead, effective earthquake mitigation strategies fall into two groups, long- and short-term. We address long-term methods first, focusing on the use of earthquake-resistant buildings. In the past, their development has been largely reactive and driven by observed failures in most recent earthquakes. As testing of building performance in future earthquakes has become more viable, there is a potential for more rapid improvements to structural design. At the same time, however, the challenges of implementation will persist, leading to a widening implementation gap between the rich and poor nations. Short-term mitigation is the topic of the final section. Over recent years, modern seismic networks have facilitated the development of rapid earthquake information systems capable of providing hazard information in the minutes after an earthquake. These systems are now beginning to provide the same information in the seconds to tens of seconds prior to ground shaking. We consider possible future applications around the world.

4.23.2 Recognizing and Quantifying the Problem

4.23.2.1 Forecasting Earthquakes at Different Spatial and Temporal Scales

The first step in seismic hazard mitigation is identification and quantification of where the hazard exists. Today, plate tectonics provides the theoretical framework for identifying and characterizing seismic source regions: where earthquakes have occurred in the past, earthquakes will occur in the future. But before the development of plate tectonic theory in the late 1960s, the same concept was in use to forecast future earthquakes. In a letter to the *Salt Lake Tribune* in 1883, G. K. Gilbert reported the findings of his field work along the Wasatch Front. He noted that the fault scarps were continuous along the base of the Wasatch with the exception of the segment adjacent to Salt Lake City where a scarp was missing. He concluded that there had been no recent earthquake on the section adjacent to the city, and this section was therefore closer to failure. In his study of deformation associated with the 1906 San Francisco earthquake, H. F. Reid built on Gilbert's model to develop elastic rebound theory that remains the basis of our understanding of the earthquake cycle today (Reid, 1910). In the elastic rebound model, the relative motion between two adjacent tectonic plates is accommodated by elastic deformation in a wide swath across the plate boundary. Once the stress on the plate boundary fault exceeds the strength of the fault, rupture occurs and the accumulated deformation across the plate boundary collapses onto the fault plane.

This cyclicity to earthquake rupture is the basis of the seismic-gap method of earthquake forecasting. If a fault segment fails in a quasiperiodic series of characteristic earthquakes, then the recurrence interval between events can be either estimated from the dates of past earthquakes or calculated by taking the characteristic slip during an earthquake and by dividing the long-term slip rate of the fault. Reported successes of seismic-gap theory include the deadly 1923 Great Kantō earthquake, the great 1946 Nankaidō earthquake and the 1944 Tōnankai earthquake (Aki, 1980; Nishenko, 1989). In 1965, Fedotov published a map showing where large-magnitude earthquakes should be expected, and his predictions were promptly satisfied by the 1968 Tokachi-Oki, 1969 Kuril Islands, and 1971 central Kamchatka earthquakes (Fedotov, 1965; Mogi, 1985). In the 1970s, the approach was applied around the globe. The estimates of relative plate motions provided by the new plate tectonic theory could be translated into slip rates across major faults. Once combined with data on the recent occurrence of large earthquakes, maps were generated identifying plate boundary segments with high, medium, and low seismic potential (Kelleher et al., 1973; McCann et al., 1979).

However, the utility of the seismic-gap method for earthquake forecasting remains a topic of debate today (e.g., Jackson and Kagan, 1993; Kagan and Jackson, 1991, 1995; Nishenko, 1989, 1991; Nishenko and Sykes, 1993). Challenges to its practical application include the incomplete historic record of earthquakes, making it difficult to estimate recurrence intervals, and difficulty in identifying the characteristic earthquake for a given fault segment. Earthquakes are also observed to cluster in space and time. Mogi (1985) proposed that plate boundary segments go through alternating periods of high and low activity, and the earthquake catalog suggests alternating periods of subduction versus strike-slip earthquake activity (Romanowicz, 1993). Laboratory experiments of stick-slip behavior show that rupture occurs at irregular intervals with variable stress drops. This implies that the state of stress before and/or after each earthquake is also variable. In Reid's original development of elastic rebound theory, he forecast that the next earthquake should be expected when "the surface has

been strained through an angle of 1/2000″ (Reid, 1910). However, he also points out that this assumes a complete stress drop, that is, the release of all accumulated strain, by the 1906 San Francisco earthquake.

The Parkfield earthquake prediction experiment is one of the more famous applications of seismic-gap theory. Three *M* 6 earthquakes located close to Parkfield in central California were instrumentally recorded in 1922, 1934, and 1966. Other data suggest an additional three events in 1857, 1881, and 1901 with a similar size and location. The similar recurrence interval of 22 years for the six events; the similar waveforms for the 1922, 1934, and 1966 events; and similar foreshock patterns prior to 1934 and 1966 make this one of the strongest cases for a characteristic earthquake (Bakun and McEvilly, 1979). Based on this evidence, Bakun and Lindh (1985) predicted that the next earthquake was due in 1988 with a 95% confidence that it would occur before 1993. An *M* 6.0 earthquake did occur on the Parkfield segment of the San Andreas, but not until 28 September 2004. While it had the same magnitude as previous events, the characteristics of its rupture were different (e.g., Langbein et al., 2005).

These examples show that while the concept of recurring seismicity is useful for forecasting future seismic hazard, the application of a recurrence interval to predict the timing of the next earthquake is fraught with uncertainties. When viewed as a stationary series, past earthquake history can be used to forecast the probability of an earthquake over long time periods (hundreds of years), and this forms the basis of the probabilistic seismic hazard analysis (PSHA) discussed in the next section. However, as the spatial and temporal scales for the forecast become smaller, the uncertainties in those forecasts become greater. The challenge is to provide forecasts that are considered relevant by society, a society that at best plans for time periods of years to decades.

4.23.2.2 Global Seismic Hazard

The United Nations designated the 1990s the International Decade for Natural Disaster Reduction. The Global Seismic Hazard Assessment Program (GSHAP) was part of this effort and had the goal of improving global standards in seismic hazard assessment (Giardini, 1999). From 1992 to 1998, an international collaboration of scientists conducted coordinated probabilistic seismic hazard analyses on a regional basis and combined them into the uniform global seismic hazard map shown in **Figure 1** (Giardini et al., 1999). The maps present the levels of peak ground acceleration (PGA) with a 10% probability of exceedance (90% probability of nonexceedance) within 50 years, corresponding to a return period of 475 years. For more information on GSHAP, visit http://www.seismo.ethz.ch/GSHAP/.

PSHA was first introduced by Cornell (1968). PSHA provides the relationship between some ground motion parameter, such as PGA, and its average return interval. There are three elements to the methodology. First, the seismic sources in a region must be characterized. It is necessary to determine where earthquakes occur, how often they occur, and how large they can be. Seismicity catalogs, both instrumental and preinstrumental, form

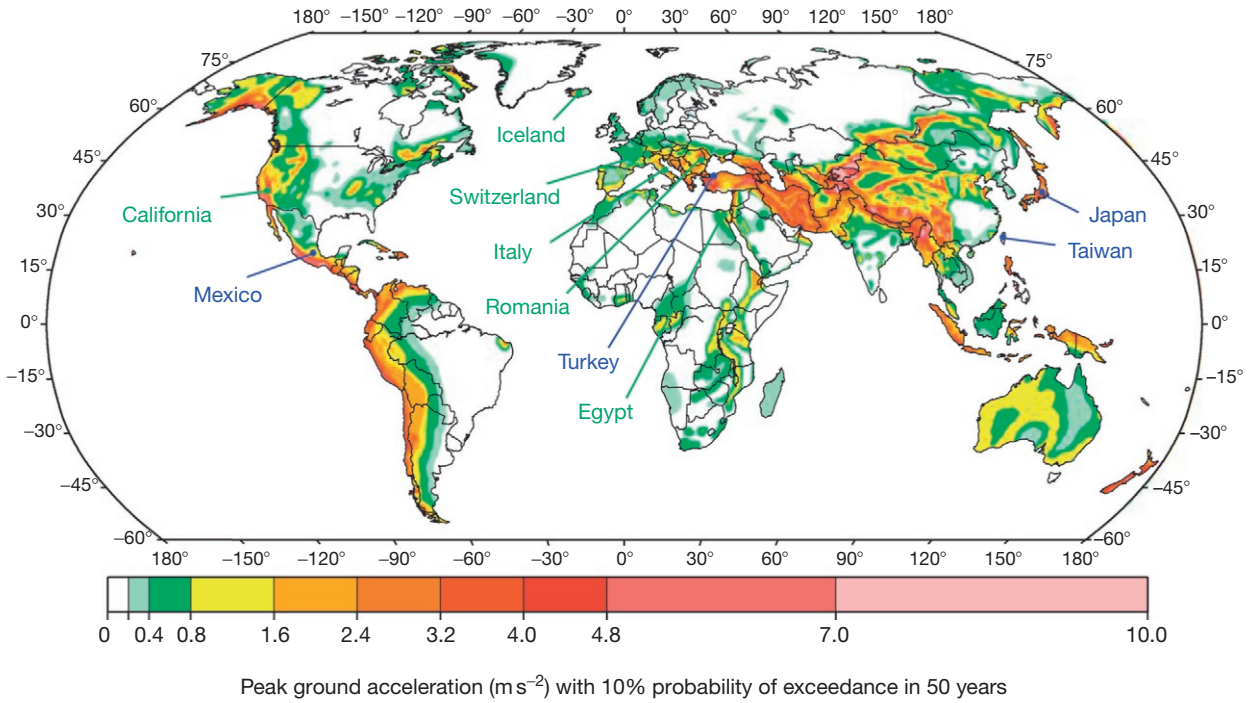


Figure 1 The global seismic hazard map developed by the Global Seismic Hazard Assessment Program (GSHAP) (Giardini, 1999). The map depicts peak ground acceleration (PGA) with a 10% probability of exceedance in 50 years, corresponding to a return interval of 475 years. The cooler colors represent lower hazard while the warmer colors are high hazard: White and green correspond to low hazard (0–0.08 m s^{-2}), yellow and light orange correspond to moderate hazard (0.08–0.24 m s^{-2}), darker orange corresponds to high hazard (0.24–0.40 m s^{-2}), and red and pink correspond to very high hazard (>0.40 m s^{-2}).

the basis of this assessment. But these catalogs are inevitably incomplete with respect to geologic timescales. Additional geodetic and geologic data are therefore included when available. Second, the expected distribution of ground shaking for all possible earthquakes is estimated. This is usually achieved using attenuation relations that describe the level of ground shaking as a function of magnitude, distance, fault type, and local site conditions. The attenuation relations are determined by regression of peak ground shaking observations for past earthquakes in the region. The quality of the attenuation relations is therefore data-limited, as we do not have observations of all possible earthquakes, particularly the larger infrequent events. For this reason, theoretical modeling of waveform propagation is now being used to improve our understanding of likely ground motions for the largest earthquakes. Finally, the probability of ground shaking at various levels is calculated by determining the annual frequency of exceedance.

To illustrate PSHA, consider the historic parameter method (McGuire, 1993; Veneziano et al., 1984). A uniform earthquake catalog is developed for the region, and attenuation functions are identified. The expected ground motion for each earthquake is then determined at every site across the region. Return periods for exceedance of ground shaking at various levels can then be tabulated and plotted to generate a hazard curve. The curve provides ground shaking level versus recurrence interval or, equivalently, probability of exceedance within some time window. The choice of ground shaking parameter varies. PGA is a short-period ground motion parameter that is proportional to force and is the most commonly mapped as the seismic provisions of current building codes specify the horizontal force a building should withstand during an earthquake. It is also the most appropriate measure for the most common building types, one- and two-story buildings, as they have short natural periods of typically 0.1–0.2 s. Other parameters that are used include peak ground velocity (PGV), which is more sensitive to longer periods and therefore appropriate for taller buildings (the natural period of buildings is typically 0.1 s per floor), and spectral response ordinates at various periods (0.3, 0.5, 1.0, 2.0 s, etc.), which are also related to the lateral forces that damage taller, longer period, buildings.

The GSHAP applied PSHA around the globe. While every effort was made to apply a uniform analysis, the differences in available data set inevitably result in some differences in the analyses for different regions (see Grunthal et al., 1999; McCue, 1999; Shedlock and Tanner, 1999; Zhang et al., 1999). Hazard curves were generated for all locations, and Figure 1 shows the PGA with a 10% probability of exceedance within 50 years. The greatest hazard is adjacent to the major transform and subduction plate boundaries: around the Pacific rim and through the broad east–west belt running from the Italian Alps to Turkey, the Zagros Mountains of Iran, the Hindu Kush, and Tian Shan and then broadening to a wider belt including the region from the Himalaya to Siberia. High seismic hazard also wraps around the coastlines of the northeast Indian Ocean, where the 2004 Sumatra–Andaman earthquake and tsunami was responsible for an estimated quarter of a million deaths. The largest recorded earthquakes are all subduction zone events; the largest three events during the last century were the 1960 Chile (M_W 9.5), 1964 Alaska (M_W 9.2), and 2004 Sumatra–Andaman (M_W 9.1) earthquakes. But the seismic hazard associated with major transform boundaries is just as large despite typically generating smaller earthquakes. This is due to the greater depth of large subduction zone earthquakes (tens of kilometers) and distance offshore, allowing attenuation of the seismic waves before they reach the land surface. By comparison, strike-slip faults rupture the shallow continental crust such as along the San Andreas Fault and the North Anatolian Fault.

4.23.2.3 Changing Seismic Risk

4.23.2.3.1 Earthquake fatalities since the 1900

The new millennium has not started well in terms of earthquake impacts on society. As of October 2006, the twenty-first century has seen almost 400 000 deaths associated with earthquakes. This represents more than 20% of the estimated 1.8 million deaths during the entire twentieth century. There is no evidence of any increase in seismic hazard; the number of earthquakes is not increasing. But is there an increase in seismic risk?

Seismic hazard analysis provides information about the likelihood of earthquakes and associated ground shaking (Figure 1). But the hazard is distinct from the 'seismic risk,' which represents the anticipated losses in a region either for a given scenario earthquake or for all anticipated earthquakes. The determination of the risk involves a convolution of the seismic hazard with population density and the properties of the built environment, including the number of buildings and the type of construction. Fragility curves are used to describe the likely damage to a building, or construction type, given different levels of ground shaking. Frequent, large earthquakes in remote areas represent high seismic hazard but low seismic risk, while moderate earthquakes directly beneath a large urban center can represent low hazard but high risk.

Figure 2 shows the cumulative number of earthquake deaths since 1900. The data come from the Significant Earthquake Database (Dunbar et al., 2006), edited to remove multiple entries and updated for the present paper. Statistical analysis of such data is notoriously difficult as it is dominated

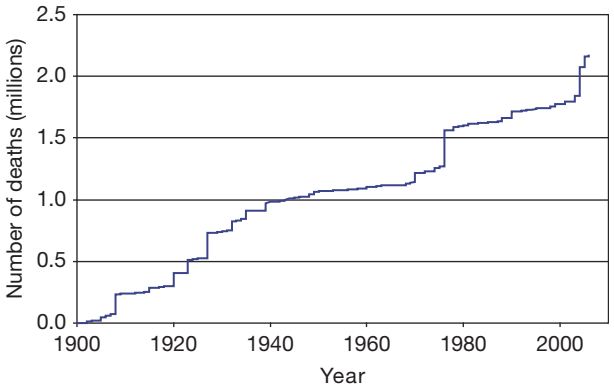


Figure 2 Cumulative number of earthquake fatalities since 1900. Note the change in character pre- and post-1940. The annual rates are ~25 000 per year pre-1940 and ~19 000 per year post-1940. Post-1940 fatalities are dominated by the Tangshan (1976) and Sumatra (2004) events with lower rates (~8000 and ~9000 per year) in between.

by infrequent high-fatality events. While there were 138 earthquakes with more than 1000 fatalities since 1900, the 10 events with the largest fatality rate caused over 60% of the deaths. The data show two trends, pre- and post-1940 (Figure 2). Prior to 1940, fatalities occur at a rate of ~25 000 per year; after 1940, the character changes and is dominated by two large fatality events (Figures 2 and 3(a)) and lower fatality rates of ~8000 per year from 1940 to 1976 and ~9000 per year from 1976 to 2004. The 1976 Tangshan earthquake is the most recent in a series of earthquakes in China with very large numbers of fatalities. The 1920 Gansu and 1927 Tsinghai earthquakes both killed an estimated 200 000; another earthquake in Gansu province killed 70 000 in 1932; and, finally, the Tangshan earthquake had an official death toll of 242 000 (as in Figure 2) but unofficial estimates as high as 655 000. The second major event in the post-1940 time series is the Sumatra earthquake and tsunami of 26 December 2004. The United Nations estimates 187 000 confirmed dead and an additional 43 000 missing. Most of these fatalities occurred in the Aceh province of Indonesia, at the northern end of the island of Sumatra, and along the Nicobar and Andaman Islands extending to the north along the subduction zone. Sri Lanka to the west and Thailand to the east were also heavily affected. The fatalities from this event are therefore more distributed than the other major events since 1900 on account of the broader reach of tsunami hazard.

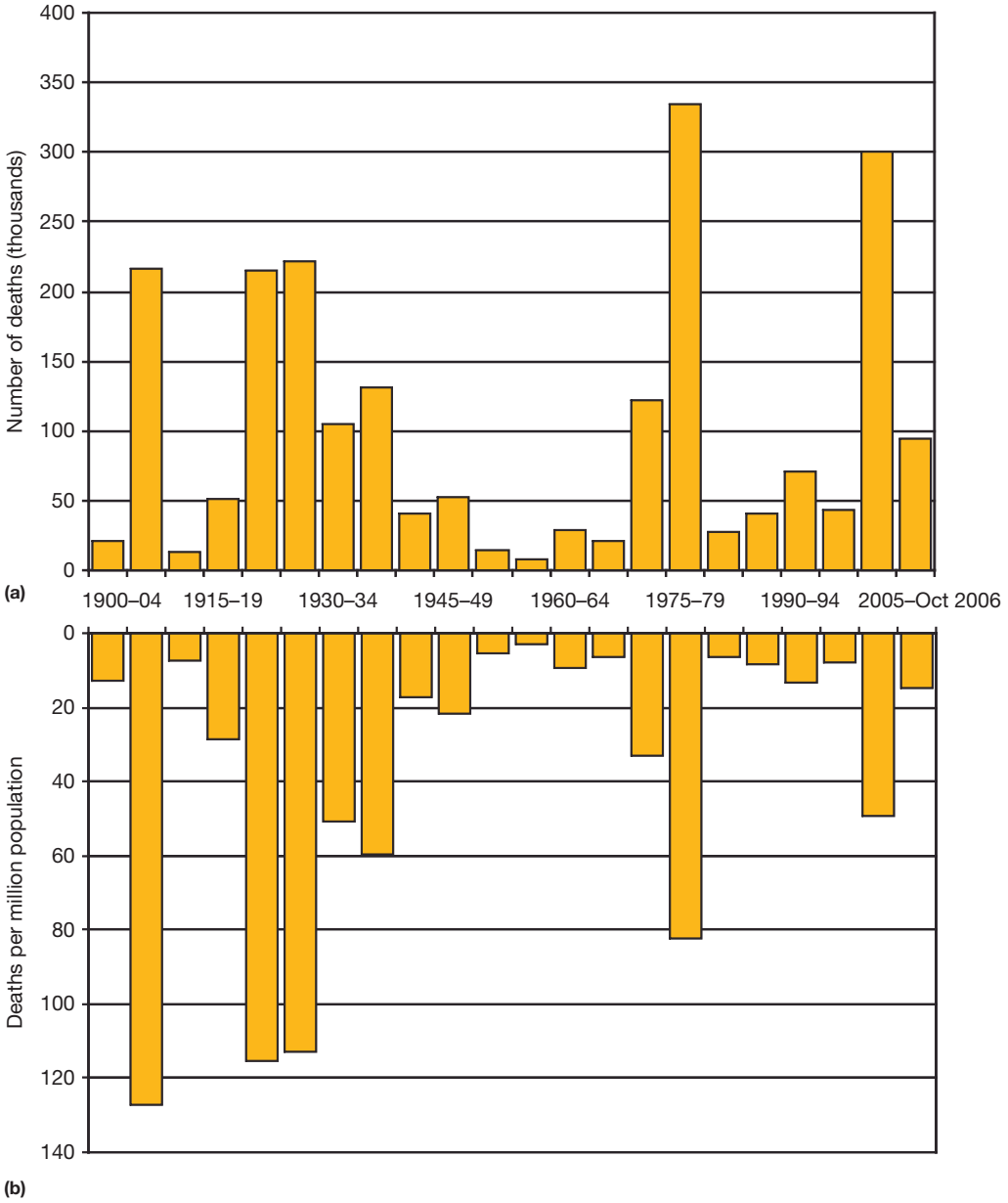


Figure 3 Earthquake-related deaths since 1900 in 5-year bins. (a) Total number of deaths. (b) Deaths per million global population. While the intervals including the 1976 Tangshan and 2004 Sumatra events have the highest number of fatalities, once normalized by global population, it is the first part of the twentieth century that has the highest rates with over 100 per million in three intervals.

The total fatality rate from 1940 to 2006 is ~19 000 per year, lower than the ~25 000 per year rate from 1900 to 1940. However, it would be a mistake to conclude that the earthquake-related fatality rate is declining as the post-1940 rate is dominated by just two events. In fact, the last five centuries of earthquake fatalities show an increasing rate. Using a best-fit power law and data from the last five centuries, Bilham (2004) estimated that the annual rate of earthquake fatalities continues to increase. While the number of deaths is increasing, it is not increasing as quickly as global population. Normalizing Bilham's best-fit annual fatality rate by global population, an individual's risk of dying in an earthquake has reduced by a factor of 2 since 1950 and a factor of 3 since 1900. This can also be seen when considering the fatalities during 5-year intervals as shown in Figure 3. The largest number of fatalities in these 5-year intervals was due to the 1976 Tangshan and 2004 Sumatra events. During these intervals, there were over 300 000 deaths, but once normalized by the global population, the highest fatality rates were during the first part of the twentieth century, when there were more than 100 deaths per million population during three 5-year intervals (Figure 3(b)).

So, are the advances in earthquake science and engineering paying off? Are we living in a more earthquake-resilient world? This conclusion would be premature for several reasons. First, the fatality rate is dominated by large impact events, and a few such events in the coming decades would reverse this trend. Second, the application of earthquake mitigation strategies is highly uneven around the globe, resulting in very different trends in regional earthquake fatality rates. Third, the distribution of global population is changing rapidly, on shorter timescales than the earthquake cycle. While the more-developed nations show zero growth, rapid growth continues in the less-developed nations, particularly in the cities. Finally, it would be irresponsible to declare success in global earthquake mitigation when the annual number of fatalities continues to increase.

4.23.2.3.2 Concentrations of risk

The high-fatality rate earthquakes recur in a relatively small number of countries. Since 1900, the 12 earthquakes causing more than 50 000 fatalities have occurred in China, Pakistan, Iran, Indonesia, Japan, Italy, and Peru. Almost half of all earthquakes causing more than 1000 deaths have occurred in these 7 countries. But the application of earthquake mitigation strategies is variable. In Japan, which has seen over 100 000 fatalities in the last century, most from the 1923 Great Kantō earthquake (also known as the 1923 Tokyo earthquake) and fire, stringent building codes are enforced, regular earthquake evacuation drills are carried out, and, most recently, an earthquake early-warning system was implemented. While in Iran, which experienced ~190 000 fatalities since 1900, the number of earthquake fatalities has tracked the population growth – one in 30 000 Iranians die in earthquakes – and the existence of earthquake building codes has had little or no effect (Berberian, 1990; Bilham, 2004).

The introduction of the medicinal control of contagious diseases at the beginning of the twentieth century finally allowed rapid growth of urban centers. Since 1950, 60% of global population growth has occurred in urban centers, and almost 50% in the lesser-developed nations (United Nations, 2004). Today, the global rural population is almost flat and the number of urban dwellers will exceed rural dwellers in 2007 for the first time. This is causing a rapid redistribution of the global population. Most of the population growth is now occurring in the less-developed nations. Within each nation, the population is migrating to the urban centers, particularly in the less-developed nations. In a series of papers, Bilham (1988, 1996, 1998, 2004) had pointed to this trend and cautioned that much higher numbers of fatalities from single events might be expected when an earthquake strikes beneath one of the growing number of large urban agglomerations.

This migration of population to the cities results in concentrations of risk. As the number of cities grows, the likelihood that an earthquake will strike a city also grows. In addition to this trend, the global distribution of the world's largest urban centers is changing. The largest cities today are in locations with a greater seismic risk than the largest cities in 1950. Figure 4 shows the seismic hazard for the world's 30 largest urban centers in 1950, 1975, 2000, and 2015. It shows that while only 10 were in regions of moderate to high hazard in 1950, this number had increased to 16 by 2000, and the trend is projected to continue. Most of the change occurred by adding new cities to the top 30 in regions of high hazard, while cities with a low hazard dropped off the list; the number of moderate hazard cities remains fairly constant. The geographic distribution of the 30 largest urban centers is shown in Figure 5. The reason for the changing hazard is clear. While the growth of cities in northern Europe and the northeastern United States has been relatively slow, rapid growth of cities in western South America and across Asia has propelled these cities with higher seismic hazard into the top 30 list.

It is tempting to associate the changing trend of global earthquake fatalities (Figure 2) with the growth of cities. Pre-1940 earthquake fatalities are more constant, while most fatalities post 1940 occurred in two events. One of the two events, 1976 Tangshan, was beneath a large city, but the fatalities in the 2004 Sumatra event were more distributed due to the

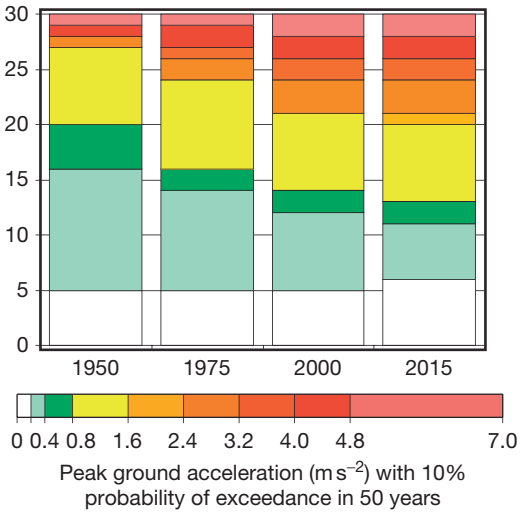


Figure 4 Seismic hazard for the 30 largest cities in 1950, 1975, 2000, and 2015 (projected). City population data from the United Nations. The seismic hazard at each city location is provided by the GSHAP map and represented as PGA with a 10% probability of exceedance in 50 years. The chart shows that cities in seismically safe regions are removed from the top 30 list as cities in hazardous regions grow more rapidly.

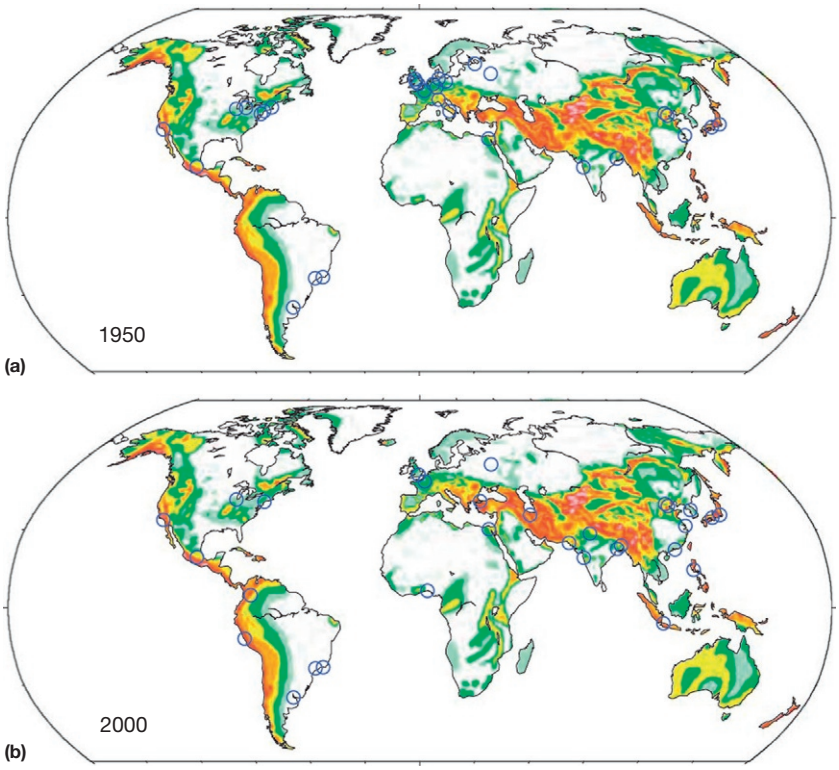


Figure 5 The locations of the 30 largest cities in (a) 1950 and (b) 2000 (blue circles) superimposed on the GSHAP map. The increased seismic hazard for the largest cities is due to relatively slow growth of cities in the eastern United States and northwest Europe while cities across Asia grow more rapidly.

tsunami. The shortness of the time history makes it impossible to be certain of the cause, and there has not yet been a large earthquake beneath a megacity.

4.23.2.4 Local Hazard and Risk: The San Francisco Bay Area

4.23.2.4.1 The San Francisco Bay Area

All seismic hazard mitigation occurs on a local scale. For this reason, it is useful to consider a case example such as the San Francisco Bay Area (SFBA). The SFBA sits within the Pacific–North American Plate boundary, which takes the form of multiple fault strands through the region (**Figure 6**). The interseismic displacement between the Pacific Plate and the western edge of the Central Valley is 38 mm year^{-1}, representing approximately 80% of the motion between the Pacific and North American Plates (d'Alessio et al., 2005). This narrow strip of land that forms the Coast Ranges of California is only ~100 km wide but has a population approaching 7 million concentrated around the bay. The SFBA has the highest density of active faults and the highest seismic moment rate per square kilometer of any urban area in the United States (WG02, 2003).

The historic earthquake record is short, believed to be complete for $M \geq 5.5$ since 1850, at which time the population exploded after gold was found in the Sierra Foothills (Bakun, 1999). Some information is available back to 1776 when the first Spanish mission, Mission Dolores, was founded. The record contains six $M \geq 6.5$ earthquakes in the SFBA in 1836, 1838, 1865, 1868, 1906, and 1989; four in the 70 years prior to 1906; and only one in the 100 years since. This change in the seismic energy release rate is believed to be due to the 'stress shadow' resulting from the 1906 San Francisco earthquake (Harris and Simpson, 1998). The 1906 event ruptured the northernmost 450 km of the San Andreas Fault from San Juan Bautista to Cape Mendocino extending through the SFBA and destroying much of San Francisco and Santa Rosa to the north. As most faults in the SFBA share a subparallel, strike-slip geometry to the San Andreas Fault, they were relaxed by the 1906 rupture.

Mapping active faults in California is the responsibility of the California Geological Survey (CGS). Under the 1972 Alquist–Priolo Earthquake Fault Zoning Act, all faults that have ruptured within the last 11 000 years are considered active, and building close to these known faults is tightly regulated to ensure that buildings are at least 50 ft from the fault trace. The CGS is now also in the process of mapping other seismic hazards including liquefaction and landslide hazards during earthquakes.

The Southern California Earthquake Center (SCEC) is a collaboration of earthquake scientists working with the goal of understanding the earthquake process and mitigating the associated hazards. While SCEC is focused on the earthquake problem in southern California, the methodologies developed by SCEC scientists to quantify earthquake probabilities and the shaking hazards associated with them are applicable everywhere, including in our chosen region of focus, the SFBA.

4.23.2.4.2 Earthquake probabilities

To evaluate the probability of future earthquakes and ground shaking in the region, the US Geological Survey established a Working Group on California Earthquake Probabilities. In several incarnations starting in 1988, the group has collected

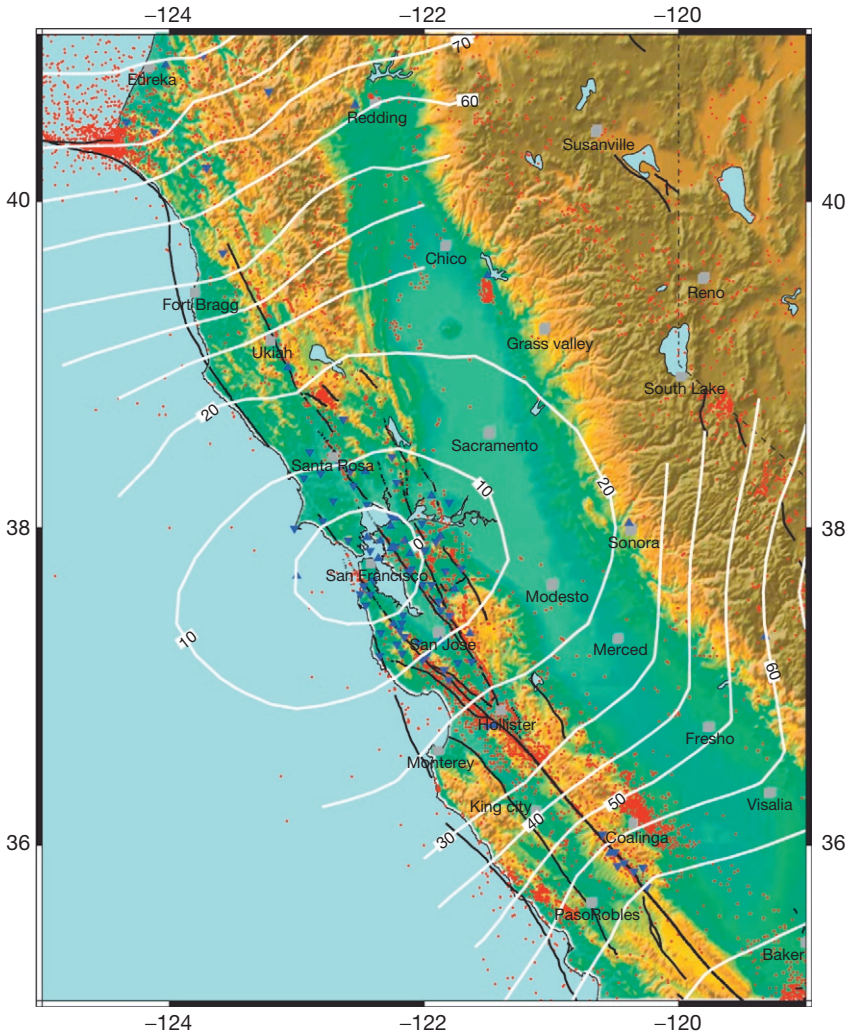


Figure 6 Map of northern California showing topography (color pallet), faults (black lines), earthquakes (red dots), seismic stations available for early warning (blue triangles and diamonds), and the warning time San Francisco could expect for earthquakes at all locations across the region (white contours, time in seconds). The warning time is estimated using ElarmS and the current seismic network and telemetry. There would be greatest warning for earthquake furthest away from the city. The existing seismic stations shown are operated by UC Berkeley and the US Geological Survey.

data and applied the most up-to-date methodologies available to estimate long-term earthquake probabilities drawing on input from a broad cross section of the Earth science community. The most recent study (hereafter WG02) was completed in 2002 (WG02, 2003). In it, the probabilities of one or more earthquakes in the SFBA, on one of the seven identified fault systems, between 2002 and 2032 were estimated. The likely intensities of ground shaking were also combined to produce a probabilistic seismic hazard map for the region, similar to the GSHAP map discussed earlier in the text. The WG02 results are shown in Figure 7.

The earthquake model used to estimate these probabilities has three elements. The first is a time-independent forecast of the average magnitudes and rates of occurrence of earthquakes on the major identified fault segments. It is derived from the past earthquake catalog. The second element includes four time-dependent models of the earthquake process to include the effects of the earthquake cycle and interactions between the fault systems. The concept of the earthquake cycle holds that after a major earthquake and associated aftershocks, another major rupture is not possible until the elastic strain has reaccumulated (Reid's elastic rebound theory). As time goes by, the probability of an earthquake therefore increases. A major earthquake also reduces the stress on any adjacent faults with a similar orientation, generating a stress shadow. This has been observed both in numerical models and in the reduced seismicity on faults adjacent to the San Andreas after the 1906 rupture. In the SFBA, both the 1906 event and the more recent 1989 Loma Prieta earthquake cast stress shadows. The third element of the earthquake model characterizes the rates of background seismicity, that is, earthquakes that do not occur on the seven major fault systems. The 1989 Loma Prieta event was one such earthquake. These various earthquake models provide different estimates of earthquake probabilities. WG02 uses expert opinion to determine the relative weight for each probability estimate derived from each model.

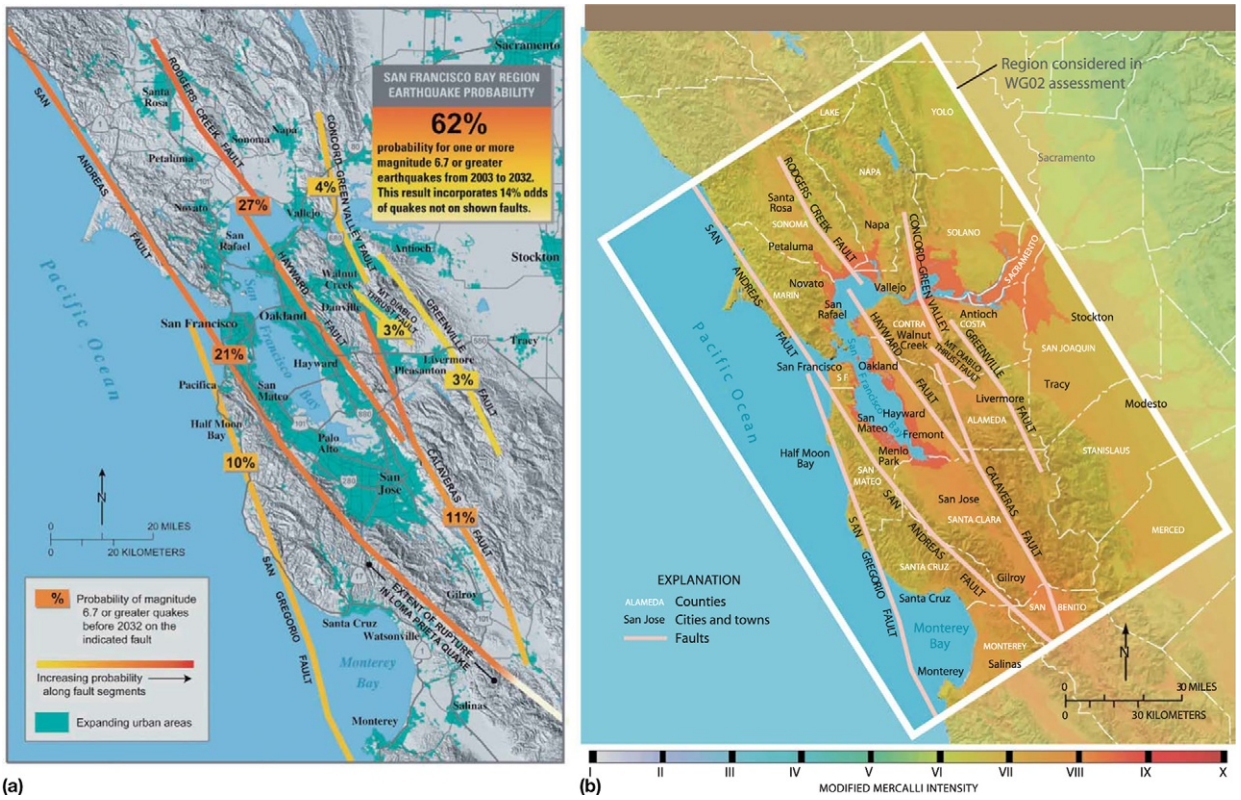


Figure 7 (a) Map of the San Francisco Bay Area (SFBA) showing the urban areas and the probabilities of $M \geq 6.7$ earthquakes by 2032. The probability of such an event in the SFBA is 62%; the probabilities of an $M \geq 6.7$ earthquake on each fault are also indicated. (b) Shaking intensities with a 50% probability of exceedance by 2032. The soft sediments and landfill around the bay and delta are where the shaking hazard is the greatest. Both figures are taken from USGS Fact Sheet 039-03 (2003).

Figure 7 shows the WG02 results. It is estimated that there is a 62% probability of one or more $M \geq 6.7$ earthquakes in the SFBA from 2002 to 2032. As shown in Figure 7(a), the probability of one or more $M \geq 6.7$ events is greatest on the Hayward–Rodgers Creek and San Andreas Faults, which have probabilities of 27% and 21%, respectively. The estimated uncertainties in these numbers are substantial. For the SFBA, the 95% confidence bounds are 37–87%. For the Hayward–Rodgers Creek and San Andreas Faults, the bounds are 10–58% and 2–45%, respectively. A critical source of this uncertainty is the extent to which the SFBA has emerged from the stress shadow of the 1906 San Francisco earthquake. Simple elastic interaction models suggest that the region should have emerged from the stress shadow, while the low seismicity rates for the last century would suggest that the SFBA remains within the shadow. Rheological models of the crust and uppermost mantle, and perhaps the 1989 Loma Prieta earthquake, suggest that the region may just be emerging now. If the region is emerging, then it can expect an increase in the number of major events over the next few decades.

4.23.2.4.3 Future losses

Just as the world has not experienced a major earthquake beneath a megacity, the United States has not experienced a major earthquake directly beneath one of its cities. The two most damaging earthquakes were the 1989 Loma Prieta earthquake (which was beneath the rugged mountains 100 km south of San Francisco and Oakland) and the 1994 Northridge earthquake (which, although centered beneath the populated San Fernando Valley, caused the strongest ground shaking in the sparsely populated Santa Susana Mountains to the north). Each event caused ~60 deaths and the estimated damages were $10 and $46 billion for Loma Prieta and Northridge, respectively (in 2000 US dollars). While the impacts were significant, the events were relatively moderate in damage.

The ground shaking estimates, such as those that are part of WG02, provide the basis for loss estimation. Loss estimation methodologies use data on the locations and types of buildings, ground shaking maps for scenario or past earthquakes, and fragility curves relating the extent of damage to the ground shaking for each building type, to estimate the total damage from the event. The worst-case scenario considered for northern California is a repeat of the 1906 San Francisco earthquake. The losses have been estimated at $170–225 billion for all related losses including secondary fires and toxic releases (RMS, 1995), a factor of 2 greater than the $90–120 billion loss estimate for property alone (Kircher et al., 2006). It is estimated that the number of deaths could range from 800 to 3400 depending on the time of day, and 160 000–250 000 households will be displaced (Kircher et al., 2006). An earthquake rupturing the length of the Hayward–Rodgers Creek Fault is estimated to cause $40 billion in damage to buildings alone (Rowshandel, 2006).

These estimates of seismic hazard and risk provide a quantitative basis for earthquake hazard mitigation in the region.

The choice of a relatively short, 30-year time window by WG02 has the advantage that it is a similar timescale to that of property ownership. But, as pointed out earlier in the text and by WG02, the reliability of PSHA decreases as the temporal and spatial scales decrease. Our observations of large ($M > 6.5$) earthquakes in California are limited. Many of the recent damaging earthquakes occurred on faults that had not been recognized, including the two most damaging earthquakes, the 1989 Loma Prieta and 1994 Northridge events. While 'background seismicity' is included in the seismic hazard estimates, these events are a reminder of the limitations to our current understanding of the earthquake hazard. These hazard and risk estimates are therefore most appropriately used to motivate broad efforts to mitigate seismic hazard across the entire SFBA rather than efforts along a specific fault segment. The limitations in our observational data set also caution against becoming too 'tuned' in mitigation strategy. The use of multiple mitigation strategies will prevent overreliance on a single, and possibly limited, model of future earthquake effects.

4.23.3 The 'Holy Grail' of Seismology: Earthquake Prediction

"When is the big one?" is the first question asked by every member of the public or press when they visit the Berkeley Seismological Laboratory. Answering this question, predicting an earthquake, is often referred to as the holy grail of seismology. In this context, a prediction means anticipating the time, place, and magnitude of a large earthquake within a narrow window and with a high enough probability that preparations for its effects can be undertaken (Allen, 1976). For the general public, answering this question is the primary responsibility of the seismological community.

The public considers earthquake prediction important because it would allow evacuation of cities and prevention of injury and loss of life in damaged and collapsed buildings. However, the seismology and engineering communities have already developed a strategy to prevent building collapse by identifying the likely levels of ground shaking and designing earthquake-resistant buildings that are unlikely to collapse. Once building codes for earthquake-resistant buildings are fully implemented, earthquake prediction would not be as important. But even before full implementation of building codes, earthquake prediction would only be partially successful as it would be capable of mitigating immediate and not long-term impacts of earthquakes. A prediction would allow for evacuations, but the ensuing earthquake would leave the urban area uninhabitable and only a fraction of the prior occupants would likely return.

It is possible to make high-probability short-term predictions for hurricanes as was done in the case of Hurricane Katrina in August 2005. Still, an estimated 1800 people were killed when New Orleans and other areas of Louisiana and Mississippi were inundated by flood waters. In New Orleans, 80% of the city was flooded, destroying much of the housing and infrastructure, and it is not yet clear what proportion will be replaced. One year later, the population of New Orleans was less than half its pre-Katrina level and roughly equivalent to what it was in 1880. If the built environment was designed to withstand a hurricane of Katrina's strength, these lives would not have been lost and New Orleans would still be thriving.

For the scientific community, earthquake prediction has a much broader meaning, encompassing the physics of the earthquake process at all timescales. The long-term probabilistic forecasts described in the previous section are predictions, but they have low probabilities of occurrence over large time windows. There is currently no approach that has consistently predicted large-magnitude earthquakes, and most seismologists do not expect such short-term predictions in the foreseeable future. While many advances have been made in understanding crustal deformation, stress accumulation, rupture dynamics, friction and constitutive relations, fault interactions, and linear dynamics, a lack of understanding of the underlying physics and difficulty in making detailed field observations mapping the spatial and temporal variations in structure, strain, and fault properties make accurate short-term predictions difficult.

In addition to these observational constraints, earthquakes are part of a complex process in which distinct structures such as faults interact with the diffuse heterogeneity of the Earth's crust and mantle at all scales. Even simple mechanical models of the earthquake process show chaotic behavior (Burridge and Knopoff, 1967; Otsuka, 1972; Turcotte, 1992), suggesting it will be difficult to predict earthquakes in a deterministic way. Instead, it may only be possible to make predictions in a statistical sense with considerable uncertainty (Turcotte, 1992). Kanamori (2003) detailed the important sources of uncertainty: (1) the stress accumulation due to relatively constant plate motion can be modified locally by proximal earthquakes; (2) the strength of the seismogenic zone may change with time, say, due to the migration of fluids; (3) predicting the size of an earthquake may be difficult depending on whether a small earthquake triggers a large one; and (4) external forces may trigger events as observed in geothermal areas after large earthquakes.

Despite these challenges, the search for the silver bullet – an earthquake precursor – continues. As pointed out by Kanamori (2003), there are two types of precursors. For the purpose of short-term earthquake prediction, the identification of a single precursor before *all* large-magnitude events is desirable. To date, no such precursor has been identified as far as we know. However, unusual precursory signals have been observed before one or perhaps a few earthquakes. These precursors may be observed before future earthquakes and are therefore worthy of research effort. The list of observed precursors includes increased seismicity and strain, changes in seismic velocities, electrical resistivity and potential, radio-frequency emission, and changes in ground water levels and chemistry (see Rikitake, 1986).

The one successful prediction of a major earthquake was prior to the 1975 Ms 7.3 Haicheng (China) event. More than 1 million people lived near the epicenter, and a recent evaluation of declassified documents concludes that an evacuation ordered by a local county government saved thousands of lives (Wang et al., 2006). There were two official middle-term predictions (1–2 years). On the day of the earthquake, various actions taken by provincial scientists and government officials constituted an imminent prediction, although there was no official short-term (a few months) prediction. A foreshock

sequence consisting of several hundred events triggered the imminent prediction; other precursors including geodetic deformation; changes in groundwater level, chemistry, and color; and peculiar animal behavior are also reported to have played a role (Wang et al., 2006). What is not known is how many false predictions were made prior to the evacuation, nor is it known how many earthquake evacuation orders have been made across China. The initial euphoria over the successful evacuation was soon dampened by the Tangshan earthquake the following year for which there was no prediction.

Extensive literature exists detailing the specifics of the various proposed earthquake prediction methodologies and other reported cases of earthquake prediction (Geller, 1996, 1997; Geller et al., 1997; Gokhberg et al., 1995; Gupta and Patwardham, 1988; Ikeya, 2004; Isikara and Vogel, 1982; Kanamori, 2003; Knopoff, 1996; Lomnitz, 1994; Mogi, 1985; Olson et al., 1989; Rikitake, 1976, 1982, 1986; Rikitake and Hamada, 2001; Sobolev, 1995; Sykes et al., 1999; Unesco, 1984; Vogel, 1979; Wyss, 1979, 1991). Expert panels are used in many countries to evaluate earthquake predictions and provide advice to governments and the public. In the United States, the National Earthquake Prediction Evaluation Council provides advice to the director of the US Geological Survey, and the California Earthquake Prediction Evaluation Council advises the governor. No short-term earthquake predictions have been made by these councils to date.

4.23.4 Long-Term Mitigation: Earthquake-Resistant Buildings

The implementation of building codes mandating the use of earthquake-resistant buildings has been highly successful in mitigating the impact of earthquakes in some regions. The number of fatalities has been reduced, and the majority of direct economic losses in recent US earthquakes (e.g., 1989 Loma Prieta, 1994 Northridge, and 2001 Nisqually) were from damage to buildings and lifelines constructed before 1976 when the Uniform Building Code was updated following the 1971 San Fernando earthquake (National Research Council, 2003). In the past, the improvement of building design was undertaken in response to observations from previous earthquakes. While improvements are still largely in response to past earthquakes today, new seismological and engineering techniques allow the development of design criteria for future likely earthquakes. Building design is also going beyond the prevention of collapse with the goal of reducing the costs of future earthquakes in addition to the number of fatalities. One of the challenges is the implementation of earthquake-resistant designs, both for new construction and for the existing building stock.

4.23.4.1 Earthquake-Resistant Design

4.23.4.1.1 Lateral forces

Following the 1891 Nobi, Japan, earthquake that killed 7000 people, John Milne laid the foundation for the building codes that were to follow (Milne and Burton, 1891). He detailed the poor performance of modern masonry construction that had recently been introduced to replace the more traditional wood construction in an effort to mitigate fires and described the great variability in damage to buildings over short distances due to the effect of soft versus hard ground. He also emphasized the need to design buildings to withstand the horizontal forces associated with earthquakes rather than just vertical forces. Similar observations were made following the 1906 San Francisco earthquake by the Lawson Commission (1908).

After the 1908 Messina–Reggio di Calabria earthquake in southern Italy, which killed 83 000 people, Panetti proposed that buildings be designed to withstand a horizontal force in proportion to their vertical load. He suggested that the first story should be able to withstand 1/12th the weight of the overlying stories and the second and third stories should be able to withstand 1/8th (Housner, 1984). In Japan, Toshikata Sano made a similar proposal. In 1915, he recommended that buildings should be able to withstand a lateral force, V, in proportion to their weight, W, such that $V = CW$, where C is the lateral force coefficient expressed as a percentage of gravitational acceleration. But it was not until the 1923 Great Kantō earthquake, which killed 100 000 people, that Sano's criteria became part of the Japanese Urban Building Law Enforcement Regulations released in 1924 (Whittaker et al., 1998). In the Japanese regulations, C was set at 10%g. Following the 1925 Santa Barbara earthquake in the United States, several communities adopted Sano's criteria with $C = 20\%g$. Sano's recommendation was also adopted in the first release of the US Uniform Building Code in 1927, where the value of C was dependent on the soil conditions (National Research Council, 2002).

4.23.4.1.2 Strong-motion observations

While building codes were mandating earthquake-resistant designs as early as the 1920s, there were still no instrumental observations of the actual ground motions responsible for building damage. Milne and colleagues designed and built the first effective seismographs in Japan in the late 1880s. The first instruments in the United States were installed at the Lick Observatory of UC Berkeley in 1887 (Lawson, 1908). By the 1920s, seismological observatories had been established around the world, but they were designed to measure the weak (low-amplitude) motion resulting from distant earthquakes. It was not until the 1930s that broadband strong-motion (high-amplitude) instruments were available, capable of recording both the low-frequency shaking and high-frequency shaking responsible for the damage to buildings. The 1933 Long Beach earthquake provided the first instrumental recording in which PGAs of 29%g in the vertical and 20%g in the horizontal were observed. A larger PGA value of 33%g was observed at El Centro a few years later on an instrument 10 km from the 1940 M 7.1 Imperial Valley earthquake rupture. This remained the largest measured ground motion for 25 years, establishing the El Centro seismogram as the standard for earthquake engineering in both the United States and Japan.

Over the following decades, the strong-motion database grew, but slowly. This changed in 1971 when the M 6.6 San Fernando earthquake struck the Los Angeles region and the number of strong-motion recordings more than doubled. In this earthquake, more than 400 000 people experienced PGA in excess of 20%g, and it became clear that high-frequency PGA varied over short distances while the longer period (10 s) displacements did not (Hanks, 1975; Hudson, 1972; National Research Council, 1971). One instrument located

on the abutment of the Pacoima Dam recorded a 1 m s^{-1} velocity pulse shortly followed by a 120%*g* acceleration pulse (Boore and Zoback, 1974). The strong-motion database generated by this earthquake played an important role in the updates to the Universal Building Code, which followed in 1976. It is a testament to the importance of strong-motion networks, and the earthquake engineering research they provide for, that the majority of damage in recent US earthquakes (1989 Loma Prieta, 1994 Northridge, and 2001 Nisqually) occurred to structures built prior to the 1976 update to the Uniform Building Code (National Research Council, 2003).

Strong-motion networks continue to provide important waveform data sets for damaging earthquakes. One notable recent example was the 1999 M_W 7.6 Chi-Chi earthquake, which occurred beneath central Taiwan on 20 September 1999. The strong-motion seismic network that had recently been deployed by the Central Weather Bureau across the island provided waveforms at 441 sites, including over 60 recordings within 20 km of the fault ruptures (Lee et al., 2001). In addition to Taiwan, dense strong-motion networks with hundreds of instruments are now operational in Japan and the western United States. Many more smaller networks are operational in earthquake-prone regions around the world. They all provide crucial data when a large earthquake occurs close by, yet the infrequency of such events makes continuous funding and operation a challenge.

4.23.4.1.3 Strong-motion simulations

Advances in computational capabilities, numerical techniques, and our knowledge of the structure of fault zone regions now make it feasible to simulate earthquakes to provide estimates of likely ground motions in future events. The recent centennial of the 1906 San Francisco earthquake motivated one such study in northern California. In order to simulate ground shaking, a velocity model was first developed for northern California. The geology-based model provides three-dimensional (3-D) velocity and attenuation for the simulation using observed relationships between rock type, depth, and seismic parameters (Brocher, 2005). Seismic and geodetic data available from the 1906 San Francisco earthquake were used to map the distribution of slip in space and time on the fault plane (Song et al., 2006). Several numerical techniques were then used to simulate the earthquake rupture through the geologic model. The simulations could be calibrated by comparing the calculated peak intensities with observed intensities from the 1906 San Francisco earthquake, which were compiled into a 1906 ShakeMap (Boatwright and Bundock, 2005; Lawson, 1908). Snapshots from one of the simulations are shown in **Figure 8** (Aagaard, 2006). The peak intensities generated by the simulations reproduce the prominent features of the 1906 ShakeMap validating the simulations.

Other simulations of the 1989 Loma Prieta earthquake, for which instrumental recording of ground shaking is available, also demonstrate that the simulations replicate the amplitude and duration of the observed shaking at frequencies less than 0.5 Hz (Aagaard, 2006; Dolenc et al., 2006). Given likely slip distributions of future earthquakes, these simulations can now provide estimates of the ground shaking in the form of complete seismic waveforms. The results of one study on the southern San Andreas Fault are shown in **Figure 9** (Olsen et al., 2006). The source rupture is along the San Bernardino Mountains and Coachella Valley segments, which are considered more likely to rupture in the coming decades as they have not ruptured since 1812 and 1960. The slip distribution of the 2002 M_W 7.9 Denali, Alaska, earthquake was used for the rupture after scaling it for an M 7.7 rupture. The velocity structure was provided by the SCEC Community Velocity Model (Kohler et al., 2003), and ground shaking is calculated for frequencies of 0–0.5 Hz just as in the northern California simulations. When the rupture propagates from the southeast to the northwest, the directivity effect produces large amplitude ground motions in the Los Angeles metropolitan region. When the fault rupture is to the east of Los Angeles, the chain of sedimentary basins (the San Bernardino, Chino, San Gabriel, and Los Angeles Basins) running westward from the northern termination of the rupture funnels seismic energy toward the downtown. The seismograms superimposed on **Figure 9** show velocities of more than 3 m s^{-1}. When the rupture propagates to the southeast, the ground shaking in LA is an order of magnitude smaller (Olsen et al., 2006).

These simulations are providing new insights into seismic wave propagation and help identify the geologic structures that control strong ground shaking. The uncertainties in the predicted ground shaking result from limitations in the velocity models, the numerical techniques, and the unknown future slip distributions. However, these simulations allow us to explore the range of possible ground motions that we might expect for earthquake ruptures that are evident in the geologic record but not the historic or instrumental records.

4.23.4.1.4 New seismic-resistant designs

As strong-motion waveforms became available to the engineering community, the complexity of surface ground motions and their interaction with buildings became apparent. Rather than containing a dominant period, the seismic waveforms were found to be more like white noise over a limited frequency range. Housner et al. (1953) proposed to reduce waveforms to a response spectrum, which is the maximum response of single degree-of-freedom oscillators with different natural periods and (typically) 5% internal damping to a recorded waveform. When the response is multiplied by the effective mass of a building, it constrains the lateral force the building would experience and should therefore be able to sustain.

The response spectra are still widely used today, but numerical techniques now allow for much more complex nonlinear modeling of buildings during ground shaking. Such modeling allows testing of new seismic-resistant designs using past earthquake recordings and future earthquake scenarios. The 1994 M_W 6.7 Northridge earthquake in southern California exposed a vulnerability in steel moment frame buildings. Moment frames resist the lateral forces in an earthquake through bending in the rigidly connected beams and columns. Due to construction practices and the use of nonductile welds, a substantial number of connections fractured in the earthquake. The Universal Building Code was updated accordingly in 1997 (UBC97). But the question remains as to how these building will behave in a future larger magnitude earthquake. Krishnan et al. (2006) explored this question using a numerical simulation of two M_W 7.9 earthquakes on the section of the San Andreas that last ruptured in 1857. They first calculated synthetic waveforms at

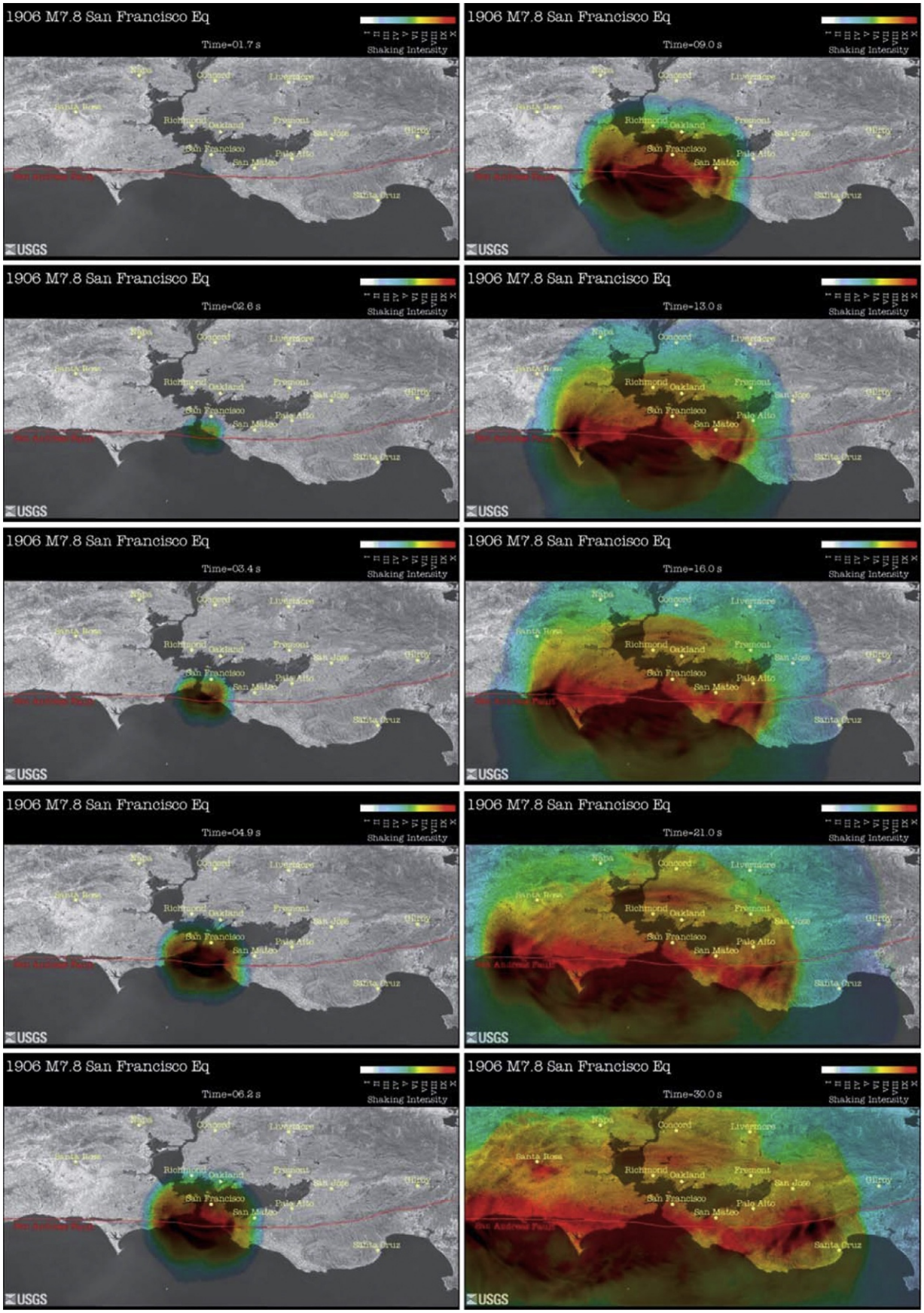


Figure 8 Simulation of the 1906 rupture along the San Andreas Fault. Each map shows the SFBA, north is to the left, and the San Andreas Fault is in red. The sequence of snapshots shows the peak ground shaking intensity (MMI) 1.7, 2.6, 3.4, 4.9, 6.2, 9.0, 13.0, 16.0, 21.0, and 30.0 s after the rupture initiates. Figures provided by Aagaard (2006). See http://earthquake.usgs.gov/regional/nca/1906/simulations/.

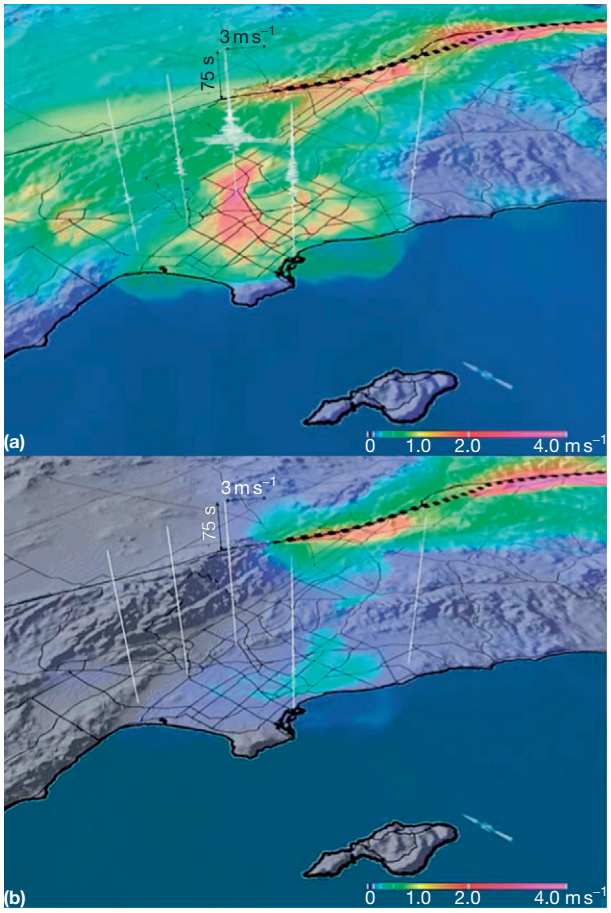


Figure 9 Simulation of an *M* 7.7 rupture on the southernmost segments of the San Andreas Fault. The section of the fault to rupture is shown by the string of black squares. The color pallet shows the peak ground velocity (PGV). (a) Rupture from the south to the north showing the funneling of energy toward the Los Angeles Basin west of the rupture. (b) When rupturing from north to south, the amplitudes in the Los Angeles Basin are an order of magnitude smaller. Reproduced from Olsen KB, Day SM, and Minster JB (2006) Strong shaking in Los Angeles expected from southern San Andreas earthquake. *Geophysical Research Letters* 33: L07305.

various locations across southern California and then simulated the effect of the ground shaking at each location on two 18-story steel moment frame buildings, one based on pre-UBC97 code and one that was post-UBC97. Krishnan et al. concluded that if the rupture propagated north to south (toward Los Angeles), then the pre-UBC97 building would likely collapse in the San Fernando Valley, Santa Monica, West Los Angeles, Baldwin Park, Compton, and Seal Beach. The post-UBC97 building would likely survive in most locations except the San Fernando Valley. This type of modeling is currently confined to the academic community; however, there is the potential to bring the lessons learned to bear on future construction practices.

Building codes for most buildings are currently focused on 'life safety,' the prevention of fatalities in an earthquake. Fatalities mostly occur due to building collapse. The goal of building codes is therefore to prevent collapse in order to get everyone out alive. With a few exceptions, codes are not intended to keep buildings in service after an earthquake and a building that performed 'well' may still need to be demolished. Earthquake engineering is now looking beyond life safety to further reduce the damage to a building at specific levels of ground shaking. Performance-based seismic design (PBSD) is one approach that focuses on what to achieve rather than what to do. The implementation of PBSD concepts will therefore lead to buildings that combine the current prescriptive building codes to prevent collapse with owner-selected design components to reduce the damage to economically acceptable levels. As a result, we can expect not only reduced fatalities in future earthquakes but also reduced economic losses that would be a reversal of the current trend of increasing economic losses (National Research Council, 2003). This poses challenges for both the seismological and engineering communities. While it is the low-frequency energy that is responsible for damage to buildings, damage to the building content is more sensitive to higher frequencies, greater than the frequency content of current ground motion simulations. For the engineering community, PBDS requires much more detailed knowledge of the performance of building components than the current prescriptive methods.

Building code requirements for critical facilities such as nuclear power plants, dams, hospitals, bridges, and pipelines are usually greater than the life safety standard currently used for homes and offices. The design criteria are continued operation for safety reasons, for example, dams and nuclear power plants, and provision of recovery services in the aftermath of an earthquake, for example, hospitals. The engineering of these facilities is usually site-specific. One example of successful engineering of a critical facility is the Trans-Alaska Pipeline System, a 48 in. diameter pipeline carrying over 2 million barrels of North Slope oil to the Valdez Marine Terminal every day. The pipeline crosses three active fault traces and was designed to withstand the maximum credible ground shaking and displacements associated with each. One of the intersected faults is the Denali Fault, where the pipeline was designed to accommodate a right-lateral strike-slip displacement of up to 6 m by constructing the supports on horizontal runners. The 3 November 2002 M_W 7.9 Denali earthquake ruptured over 300 km of the Denali, Totschunda, and Susitna Glacier faults, including the section beneath the pipeline. The displacement at the pipeline was 5.5 m, and there was only minor damage to some of the supports that had been displaced several meters by the rupture (Sorensen and Meyer, 2003).

Structural control is another relatively new approach to reducing the impact of large earthquakes on various structures. The concept is to suppress the response of a building by either changing its vibration characteristics (stiffness and damping) or applying a control force. There are active, semiactive, and passive types of structural control. Active control systems are defined as those that use an external power source. The active mass damper is one such device where an auxiliary mass is driven by actuators to suppress the swaying of a building. Kajima Corporation applied this technique to its first building in 1989, and the device is capable of suppressing the response of the building to strong winds and small to medium earthquakes. The high power demand limits its effectiveness for large earthquakes. Passive systems rely on the viscoelastic, hysteretic, or other natural properties of material to reduce or dampen vibrations.

Base isolation is one example of a passive system in which large rubber pads separate a building from the ground. These pads shear during strong shaking, reducing the coupling between the building and the ground. These devices have the advantage that they require no external power, need little or no maintenance, and perform well in large earthquakes. There are now over 200 buildings around the world with base isolation systems. Finally, semiactive systems use a combination of the two approaches in that the building response is actively controlled but using a series of passive devices. Active variable stiffness and active variable damping devices are currently being used as part of semiactive systems. These semiactive systems have been installed in a few buildings in Japan as they are still in the development mode, but, as with PBSD, they hold the promise of reducing not only the number of fatalities but also the economic losses associated with future earthquakes.

4.23.4.2 The Implementation Gap

There are two implementation gaps that seriously negate the effectiveness of earthquake-resilient building design. The first is the large variability in their application or enforcement in different countries; the second is that building codes are generally only applicable to new construction.

4.23.4.2.1 The rich and the poor

Earthquake-resistant design has been proven effective and building codes that include earthquake provisions have been adopted in most countries that have experienced multiple deadly earthquakes (Bilham, 2004). However, while the number of earthquake fatalities in rich countries is estimated to have decreased by a factor of 10, presumably due to better buildings and land use (Tucker, 2004), the number of fatalities in poor countries is projected to increase by a factor of 10. The 1950 M 8.6 Assam earthquake in India killed 1500 people, but it is estimated that a repeat event in the same location would kill 45 000 people (Wyss, 2004), an increase by a factor of 30 in a region where the population has increased by a factor of 3. Similarly, a repeat of the 1987 M 8.3 Shillong earthquake would kill an estimated 60 times as many people as in 1987 (Wyss, 2004). During that period, the population has increased by a factor of 8, again suggesting an order of magnitude increase in the lethality of earthquakes. This increase is largely due to the replacement of single-story bamboo homes with multistory, poorly constructed, concrete frame structures, often on steep slopes (Tucker, 2004).

Berberian (1990) investigated the earthquake history in Iran. He concluded that the adoption of building codes has had little or no effect, largely due to lack of enforcement. The enforcement gap was also identified after the 1999 Izmit earthquake in Turkey as a major contributor to the 20 000 fatalities. Better implementation and enforcement therefore remain a priority in many earthquake-prone regions. However, the socioeconomic situation in many of these countries leaves earthquake risk reduction low on the priority list of development agencies. Most aid organizations continue to operate in a response mode to natural disasters rather than a preventative one. One notable exception is GeoHazards International (http://www.geohaz.org), who are working to introduce earthquake-resistant building practices to local builders in regions of high seismic risk.

4.23.4.2.2 The new and the old

Building codes only apply to new construction. As is clear from the history of earthquake-resistant building design, every major earthquake to date has provided lessons in how not to construct buildings. Unreinforced masonry was banned for public schools in California after the 1933 Long Beach earthquake. In the most recent earthquakes, problems with moment frame buildings and the dangers of soft-story buildings were identified. After each of these earthquakes, building codes are updated. The vast majority of buildings are therefore not up to current code. Several hundred billion dollars are spent every year on construction in seismically hazardous areas of the United States. It is estimated that the additional earthquake-related requirements of building codes account for ~1% of this investment; the cost of making new buildings seismically safe is therefore small (Office of Technology Assessment, 1995). In contrast, the cost of retrofitting existing buildings is much higher, around 20% of the value of the building for most construction types. In addition to the cost, buildings usually need to be vacated during the retrofit, causing additional disruption to the occupants. One example of the retrofitting gap comes from a 2001 study of hospital seismic safety in California (Office of Statewide Health Planning and Development, 2001). The study estimated that over a third of the state's hospitals were vulnerable to collapse in a strong ($6.0 < M < 6.9$) earthquake. In Los Angeles County, more than half were vulnerable, and the ratio rises to two in three in San Francisco. The total cost of initial improvements required by state law after the 1994 Northridge earthquake totaled $12 billion; in Los Angeles County, the bill was greater than the total assessed values of all hospital property. Hospitals are considered critical infrastructure, which is why they are required to retrofit by law, but given these economic realities, the extent of the retrofits remains to be seen.

The high cost and inconvenience of retrofitting, combined with the uncertainty in the benefit, mean that few buildings are retrofitted. However, some institutions and governmental bodies have risen to the challenge. One example of an institution stepping forward to tackle this problem is UC Berkeley (Comerio et al., 2006). The university campus sits astride the Hayward Fault, considered to be one of the most hazardous faults in the SFBA. Since the university was founded, it has had a commitment to the safety of its students, faculty, and staff, and seismic-resistant designs have been used across campus. Following the 1971 San Fernando earthquake, which caused some damage to another University of California (UC) campus, weaknesses in current building practices were identified and the Universal Building Code was updated in 1976. In 1978, the UC system adopted a seismic safety policy and undertook a review of buildings across the Berkeley campus. Key buildings including University Hall, which housed the system-wide administration at the time; high-rise residence halls; and some key classroom buildings and libraries were retrofitted.

The 1989 Loma Prieta, 1994 Northridge, and 1995 Kobe earthquakes demonstrated how relatively modern buildings were still susceptible to damage during earthquakes and refocused the university on seismic safety. A complete review of

campus buildings was ordered in 1996, and it was determined that one-third of all space on campus was rated as poor or very poor, that is, susceptible to collapse in an earthquake. In 1997, the SAFER program was initiated to retrofit or replace seismically hazardous buildings across campus for life safety. The financial commitment was $20 million per year for 20 years. The most hazardous buildings were retrofitted first and the program continues today. At the same time that the SAFER program was being formulated, Mary Comerio conducted a study of the broader social and economic impacts of future earthquakes. One of the conclusions was that the campus would likely have to close for one or more semesters after an earthquake on the Hayward Fault. This posed a long-term threat to the university's existence as many students, faculty, and staff would likely move elsewhere during this period and not return. The seismic retrofit program was therefore expanded to include business continuity as a goal in addition to life safety and incorporated elements of performance-based design.

The city of Berkeley has also shown leadership in developing innovative programs to motivate the seismic retrofitting of buildings. One such program is the transfer tax incentive. On purchasing a home, one-third of the transfer tax payable to the city is available for approved seismic retrofitting of the home. This typically amounts to several thousand dollars each time a home changes hands. While an individual homeowner may not fully retrofit the home, as properties change hands over time, the building stock becomes more seismically safe. This program, in concert with other city retrofit incentives, has resulted in over 80% of single-family homes being at least partially retrofitted in the city, and an estimated 35% are fully retrofitted, making Berkeley one of the most improved cities for seismic safety in the Bay Area (Perkins, 2003).

It is even more of a challenge to motivate retrofitting of buildings that are not owner-occupied. In a program initiated in 2006, the city of Berkeley is targeting the large number of soft-story apartment buildings. Soft-story buildings have large openings in walls on the ground floor, which, as recent earthquakes have demonstrated, makes them vulnerable to collapse. The openings most commonly allow access to parking under the building or store fronts. Under the new city ordinance, soft-story buildings are first identified on a city list and owners are notified. The owner is then required to notify existing and future tenants of the earthquake hazard and postprominent seismic hazard signs. The owners are also required to have an engineering assessment of the seismic safety of the buildings and make the information available to the city. The program is designed to provide an incentive for owners to retrofit their buildings. The effectiveness of the program will depend on the extent to which tenants are concerned about seismic safety and whether there are alternative accommodations.

4.23.5 Short-Term Mitigation: Real-Time Earthquake Information

The expansion of regional seismic networks combined with the implementation of digital recording, telemetry, and processing systems provides the basis for rapid earthquake information. This process is often referred to as real-time seismology and involves the collection and analysis of seismic data during and immediately following an earthquake so that the results can be effectively used by the emergency response community and, in some cases, for early warning (Kanamori, 2005). One of the first reported calls for real-time earthquake information came in 1868 following two damaging earthquakes in SFBA in just 3 years. Following the failure of a 'magnetic indicator' for earthquakes, J. D. Cooper suggested the deployment of mechanical devices around the city to detect approaching ground motion and transmit a warning to the city using telegraph cables (Cooper, 1868). Unfortunately, his system was never implemented.

In California, the first automated notification systems provided earthquake location and magnitude information. They used the Real-Time Picker (RTP) and became operational in the mid-1980s. RTP identified seismic arrivals on single waveforms and estimated the signal duration providing constraints on earthquake location and magnitude (Allen, 1978, 1982). In the early 1990s, the systems were further developed to integrate both-short-period information and broadband information. The Caltech/USGS Broadcast of Earthquakes (CUBE) (Kanamori et al., 1991) and the Rapid Earthquake Data Integration project (Gee et al., 1996, 2003), in southern and northern California, respectively, provided location and magnitude information to users within minutes via pagers.

In Japan, real-time earthquake information systems have been developed in parallel with those in the United States. By the 1960s, single seismic stations were already being used to stop trains during earthquakes. After the 1995 Kobe earthquake, the Japanese government initiated a program to significantly increase the seismic instrumentation across the country with multiple seismic networks. The strong-motion Kyoshin network (K-Net) has over 1000 stations across the entire country with a constant station spacing of 25 km (Kinoshita, 2003). In addition, most of the ~700 short-period instruments deployed in boreholes (Hi-Net) also have strong-motion instruments at the top and bottom of the borehole (KiK-Net). Finally, a lower-density broadband seismometer network consisting of ~70 instruments with a typical station spacing of 100 km spans the entire country. These networks are operated by the National Research Institute for Earth Science and Disaster Prevention (NIED). All data are telemetered in real time and is available via the Web (http://www.bosai.go.jp). The Japan Meteorological Agency (JMA) also operates a seismic network across the country that is used for real-time earthquake information.

4.23.5.1 Ground Shaking Maps: ShakeMap and Beyond

Following the 1994 Northridge earthquake, the TriNet project (Hauksson et al., 2001; Mori et al., 1998) was designed to integrate and expand seismic networks and monitoring in southern California. In both the 1994 Northridge and 1989 Loma Prieta earthquakes, strong ground shaking occurred away from the epicenter, and there was a need to go beyond point-source information and provide better estimates of the locations of likely damage to the emergency response community. In the 1995 Kobe earthquake, it was many hours until the central government in Tokyo was aware of the full extent of damage to the city of Kobe, delaying rescue and recovery efforts (Yamakawa, 1998), again emphasizing the need for rapid automated ground shaking information after major earthquakes.

4.23.5.1.1 ShakeMap

The development and implementation of ShakeMap (Wald et al., 1999) was the response of the seismological community. The ShakeMap concept is to rapidly gather ground shaking information following an earthquake and integrate it into a map of peak ground shaking distribution. While the concept is simple, the implementation is complex, as data from different instrument types with a highly heterogeneous distribution must be integrated. The ShakeMap methodology is triggered by the identification of an earthquake, typically with $M \geq 3$, and first gathers PGA and PGV data from seismic instruments in the proximity of the earthquake. The system must wait several minutes for all stations within a few hundred kilometers to record peak ground shaking and telemeter the data to the central processing site.

Once at the central site, individual station data are first corrected for site amplification effects so they represent observations at uniform 'rock' sites. An empirical attenuation relation for an earthquake of the observed magnitude within the region is then adjusted to provide the best-fit relation for the ground shaking as a function of distance. The attenuation relation is used to generate a map of predicted rock-site ground shaking at all locations. This map is adjusted to match local station observations providing a map of ground shaking controlled by the observations close to seismic stations and the best-fit attenuation relation where there are no data. Finally, adjustments are made for site amplification effects based on mapped geology in the region. In addition to providing maps of PGA and PGV, ShakeMap also combines these data and uses scaling relations to provide estimates of instrumental Modified Mercalli Intensity (MMI) (Wald et al., 1999). MMI was developed prior to modern seismic instrumentation but still provides a useful description of the felt ground shaking and damage. More detailed information is available in the ShakeMap manual (Wald et al., 2005).

The methodology was in place for the 1999 M_W 7.1 Hector Mine earthquake providing a test of the real-time earthquake information system (Hauksson et al., 2003). A location and preliminary local magnitude estimate of 6.6 were first available 90 s after the event origin time. An energy magnitude of 7.0 was available 30 s later. These estimates were broadcast via e-mail, the Web, and the CUBE pager system within minutes. The first ShakeMap was produced within 4 min of the event. This initial map was generated using observed peak ground shaking and the best-fit attenuation relation, assuming that the ground shaking decayed as a function of distance from the epicenter. As there was only one station within 25 km of the rupture, near-fault ground shaking was estimated based on the attenuation relations. Over the following hours, ShakeMap was updated using information on the finiteness of the fault based on aftershock locations, finite-source inversions, and field observations. Broadband waveforms from more distant sites were used to model the rupture improving the estimates of near-fault shaking (Dreger and Kaverina, 2000). The final version is shown in **Figure 10**.

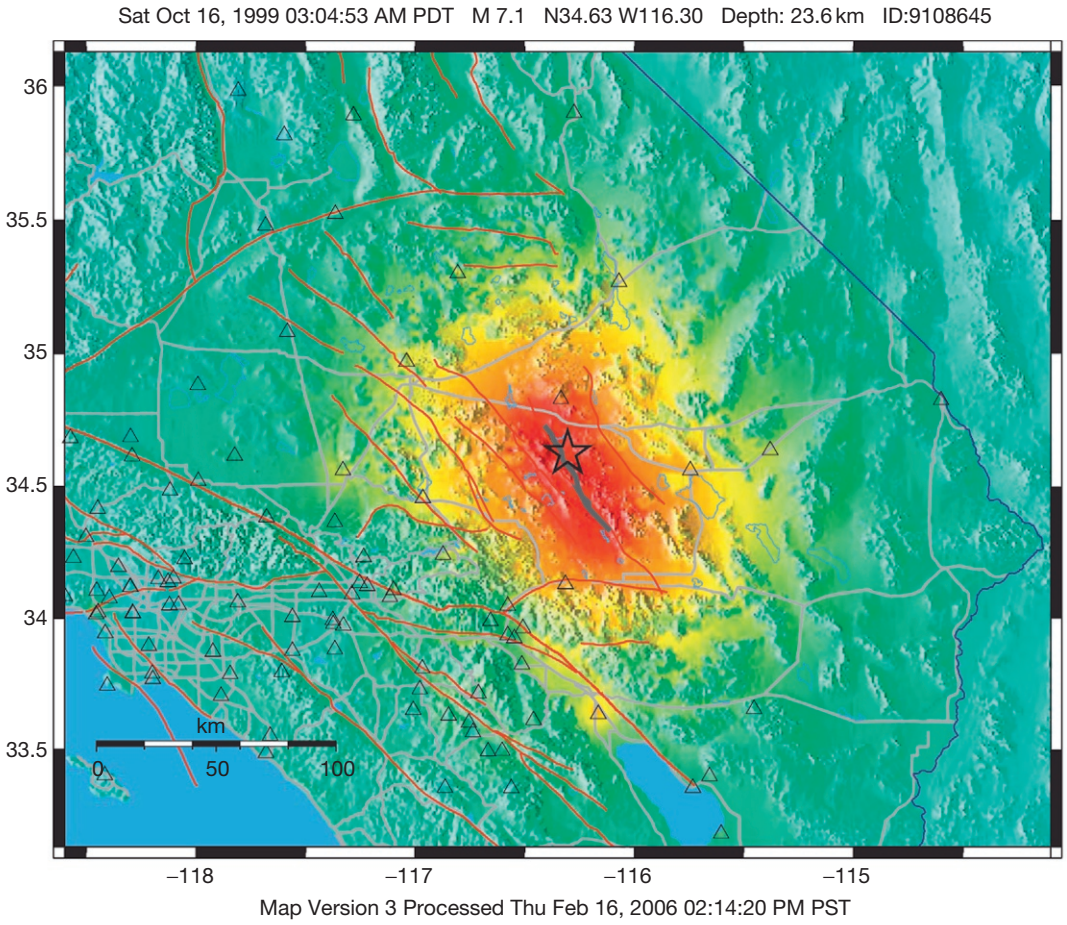
4.23.5.1.2 Rapid finite-source modeling

The ShakeMap approach works best in regions with dense station coverage. The observed ground motions then control the contouring of the maps. However, the success of ShakeMap has resulted in a desire to generate maps in regions where the station coverage is sparse to nonexistent. Broadband seismic stations can be used to model the finiteness of the source and improve the ShakeMap (e.g., Dreger and Kaverina, 2000; Ji et al., 2004). The integration of rapid and automated finite-source modeling into ShakeMap-type products represents one of the new directions in seismic hazard mitigation.

The value of finite-source information was demonstrated by the 2003 M_W 6.5 San Simeon earthquake in central California (Dreger et al., 2005; Hardebeck et al., 2004). The seismic station distribution is sparse in the region, resulting in only three observations of peak ground shaking close to the event in real time. The initial ShakeMap for the event (**Figure 11(a)**) is therefore dominated by the event location and magnitude estimate from which the radial attenuation relation is defined. In fact, the ruptured fault plane extended to the east from the hypocenter, resulting in stronger ground shaking to the east than suggested by this initial ShakeMap. **Figure 11(d)** shows the best estimate of ground shaking available today for comparison; it includes data that were not available in the initial hours after the event.

A real-time finite-source inversion scheme was developed for this scenario by Dreger and Kaverina (2000) using data from the 1992 Landers and 1994 Northridge earthquakes. Although the codes were not automated at the time of 1999 Hector Mine earthquake, they were able to use the off-line version to determine finite-source variables and forward calculate ground motions within 5 h of the event. The now-automated approach (Dreger and Kaverina, 2000) first determines a moment tensor, which typically takes 6–9 min. A series of finite-source inversions are then used to explore model space. The moment tensor provides two possible fault planes and the size of the rupture based on moment scaling relations (Somerville et al., 1999). The data are inverted for a series of line sources to test the two moment-tensor nodal planes and a range of rupture velocities. These results are available 11–20 min after the event. At this stage, the orientation and length of the fault plane can be provided to ShakeMap, allowing the ground motion to be estimated as a function of distance from the surface projection of the fault plane rather than distance from the epicenter. A 2-D inversion usually follows, providing a better description of the kinematics of the fault rupture. Finally, the kinematic model can be integrated with near-fault Green's functions to simulate near-fault waveforms, all within ~30 min of an earthquake (Dreger and Kaverina, 2000; Kaverina et al., 2002).

The first event in which the ShakeMap was rapidly updated with finite-source information was the 2003 San Simeon earthquake (Dreger et al., 2005). The earthquake occurred in a sparsely populated rural area and most of the damage occurred in the town of Paso Robles, 35 km southeast of the rupture where two people were killed. The line-source inversion was complete 8 min after the event and the 2-D inversion and predicted ground motions were available after 30 min. The ShakeMap was updated using the length and geometry of the fault plane derived from the finite source as shown in **Figure 11(b)**. The inclusion of the fault plane resulted in increased estimates of ground shaking at Paso Robles. The initial point-source ShakeMap estimated MMI of V–VI. With



Perceived shaking	Not felt	Weak	Light	Moderate	Strong	Very strong	Severe	Violent	Extreme
Potential damage	None	None	None	Very light	Light	Moderate	Moderate/heavy	Heavy	Very heavy
Peak acc. (%g)	*<.17*	*.17–1.4*	*1.4–3.9*	*3.9–9.2*	*9.2–18*	18–34	34–65	65–124	>124
Peak vel. (cm s^{-1})	<0.1	0.1–1.1	1.1–3.4	3.4–8.1	8.1–16	*16–31*	*31–60*	*60–116*	*>116*
Instrumental intensity	I	II–III	IV	V	VI	VII	VIII	IX	X+

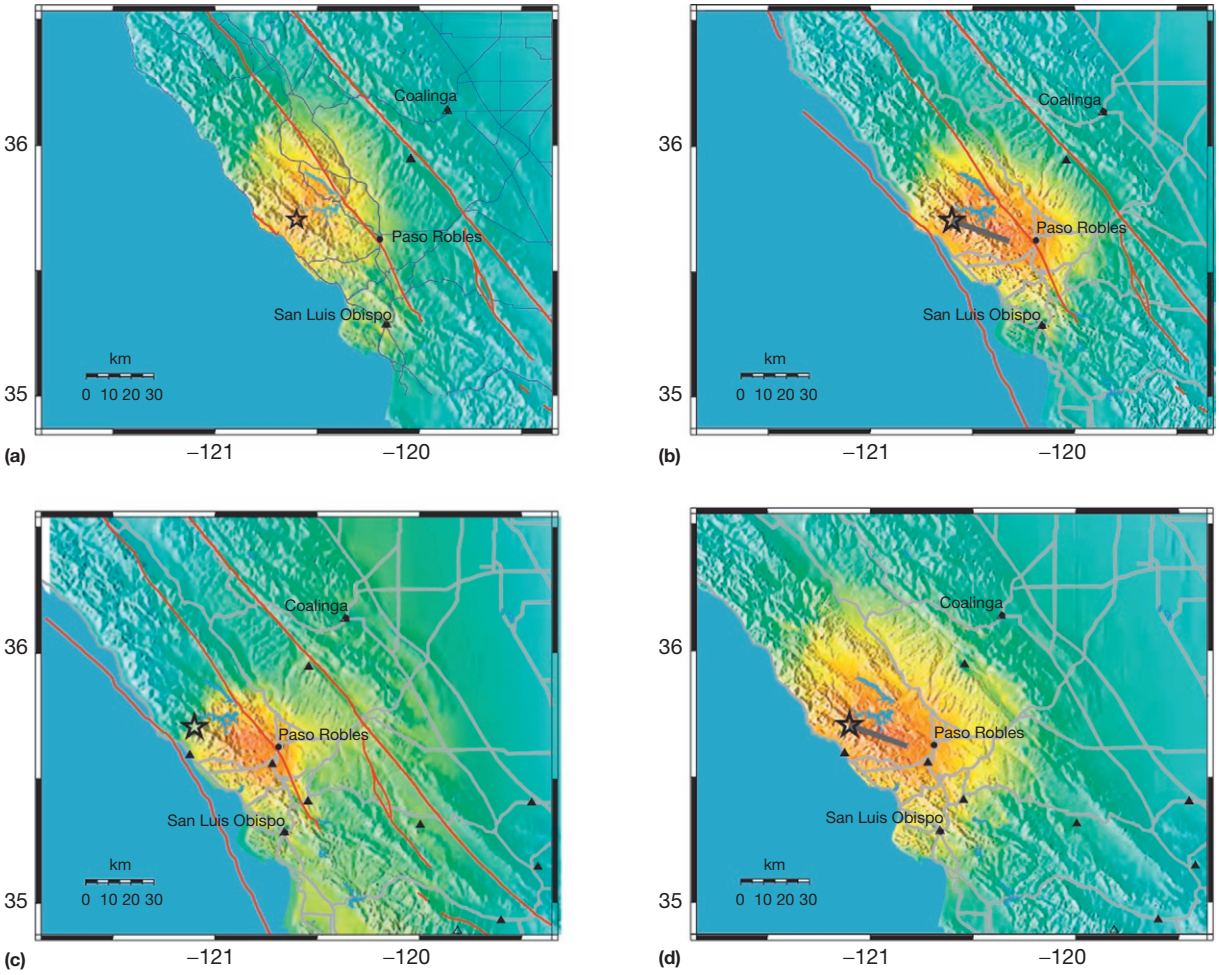
Figure 10 ShakeMap for the 1999 M_W 7.1 Hector Mine earthquake. This version includes the finiteness of the fault rupture that became available in the hours after the earthquake. There is only one seismic station within 25 km of the rupture, so attenuation relations describing MMI as a function of distance from the fault constrain the near-field MMI estimates. The star shows the epicenter and the line represents the finite extent of the fault. The color pallet shows the instrumental MMI.

the fault plane included in ShakeMap, the MMI increased to VII–VIII (compare **Figure 11(a)** and **11(b)**), which is in line with observed damage.

4.23.5.1.3 Applications of ShakeMap

Since its inception, ShakeMap has become a great success, both within the emergency response community for whom it was originally designed and with the broader public. While the 1999 Hector Mine earthquake was felt widely across the Los Angeles Basin, the ShakeMap showed that the earthquake was fairly distant, centered in the Mojave Desert (**Figure 10**). This information provided for an appropriately scaled response. One Caltrans bridge crew member reported: "I can't tell you how much time and money was saved knowing where to look [for damage]." ShakeMaps are now routinely generated in Nevada, Utah, the Pacific Northwest, and Alaska in addition to California (visit http://earthquake.usgs.gov/eqcenter/shakemap/). Other earthquake-prone regions around the world are also using and developing similar tools. The ShakeMap output also includes GIS shapefiles of ground shaking levels for use in loss estimation calculations such as HAZUS. These loss estimates are now routinely performed in the hours after moderate and large earthquakes to guide response and recovery.

ShakeMap has also become a tool for public information and education. On the day of the Hector Mine earthquake – ShakeMap's debut – more than 300 000 people visited the website. After smaller, felt earthquakes, website visits reached hundreds per second. In response to this public interest, media maps were designed with the TV audience in mind. These



Perceived shaking	Not felt	Weak	Light	Moderate	Strong	Very strong	Severe	Violent	Extreme
Potential damage	None	None	None	Very light	Light	Moderate	Moderate/heavy	Heavy	Very heavy
Peak acc. (%g)	*<.17*	*.17–1.4*	*1.4–3.9*	*3.9–9.2*	*9.2–18*	18–34	34–65	65–124	>124
Peak vel. (cm s^{-1})	<0.1	0.1–1.1	1.1–3.4	3.4–8.1	8.1–16	*16–31*	*31–60*	*60–116*	*>116*
Instrumental intensity	I	II–III	IV	V	VI	VII	VIII	IX	X+

Figure 11 ShakeMaps for the 2003 M_W 6.5 San Simeon earthquake in central California. Black triangles are seismic stations, the star is the epicenter, and the black line represents the finite extent of the fault. The color pallet shows the instrumental MMI. (a) The automated ShakeMap generated without any finite-source information. There are only three stations constraining the ground shaking estimates on the map. (b) The ShakeMap once the length and geometry of the finite source were included based on the information provided by the real-time finite-source model. (c) The ShakeMap derived from all available ground motion observation today (including those for which the waveform data had to be transported by hand on magnetic tape) but without any finite-source information. (d) The best estimate of the distribution of ground shaking intensity available today. This incorporates the finite extent of the fault and all stations in the region. Modified from Dreger DS, Gee L, Lombard P, Murray MH, and Romanowicz B (2005) Rapid finite-source analysis and near-fault strong ground motions; application to the 2003MW 6.5 San Simeon and 2004MW 6.0 Parkfield earthquakes. *Seismological Research Letters* 76: 40–48.

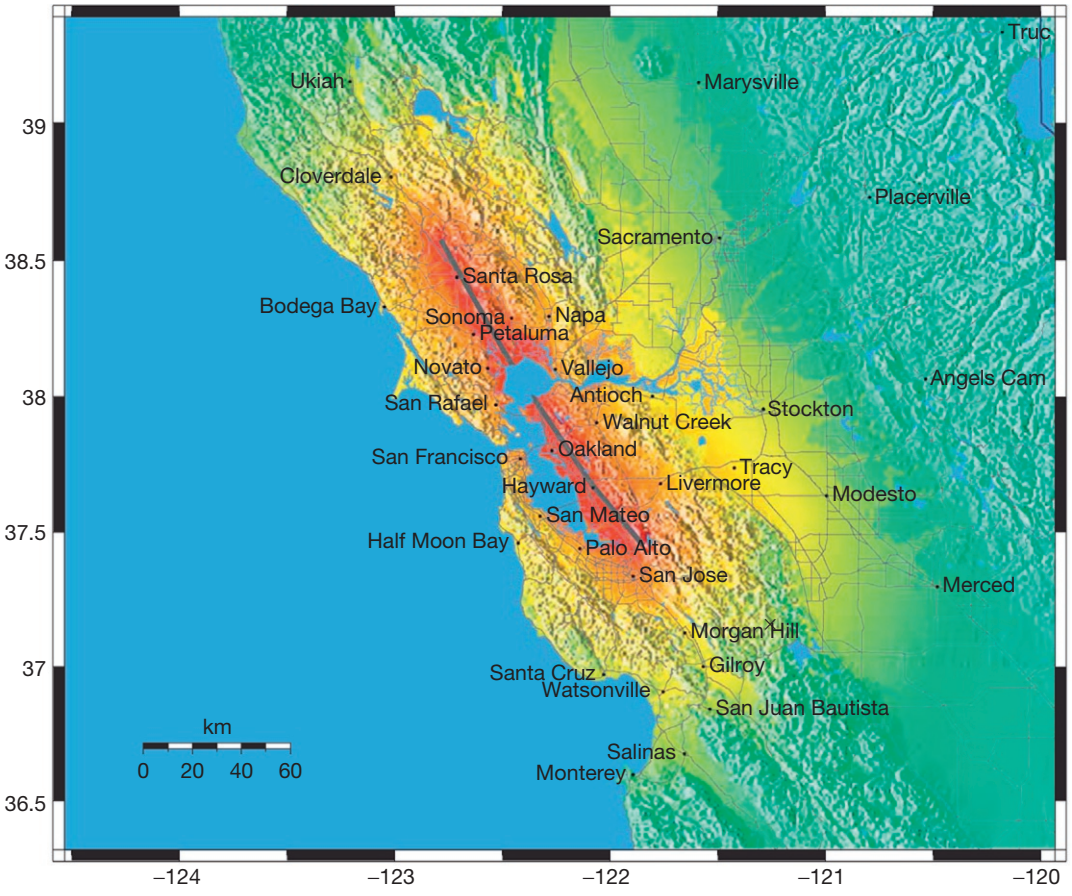
simplified versions of ShakeMap are routinely produced and often used in media coverage following earthquakes. Perhaps the best example of the public interest in the ShakeMap concept is the birth of Community Internet Intensity Maps (CIIMs), better known as "Did you feel it?" These MMI maps are generated automatically using reports of ground shaking intensity provided by the public using an Internet portal (http://earthquake.usgs.gov/eqcenter/dyfi.php). These reports are averaged by zip code and provide maps that are very similar to the instrumental MMI ShakeMaps. The CIIMs generate thousands of reports after a felt earthquake, the maximum to date was just under 30 000 after an M 5.2 near Anza, California, in June 2006 (Wald et al., 2006b). In 2004, the USGS extended the CIIM system to allow for international data collection. These ShakeMap-type products have extended the reach and the complexity of earthquake information provided to the

public. This provides an inherent educational benefit as the consumers become more informed about earthquake hazards.

The ShakeMap products for the technical user have also been expanding. Maps of the response spectral acceleration at 0.3, 1.0, and 3.0 s periods are important for estimating the effects of the shaking on particular types of buildings. This information is also available for past significant earthquakes, prior to the inception of ShakeMap, and thus provides a history of the ground shaking experienced by a particular building. These past earthquake maps are also useful for planning and training purposes in preparation for future events. Probabilistic assessments of future likely earthquakes, such as those shown for the SFBA earlier in the text, have also been used to generate scenario ShakeMaps that can be used in loss estimation and also for training. A scenario ShakeMap for a rupture of the Hayward–Rodgers Creek Fault is shown in Figure 12. Finally, ShakeCast is a new mechanism for the delivery of ShakeMap that can also be used to trigger user-specific postearthquake response protocols. For example, utilities, transportation agencies, and other large organizations can automatically determine the shaking at their facilities, set thresholds for notification, and notify responsible staff when appropriate. More information on the range of rapid postearthquake information products provided by the USGS is available online at http://earthquake.usgs.gov/.

4.23.5.1.4 Global earthquake impact: PAGER

All of the rapid postearthquake information discussed in the preceding text is seismic hazard information. However, it is the

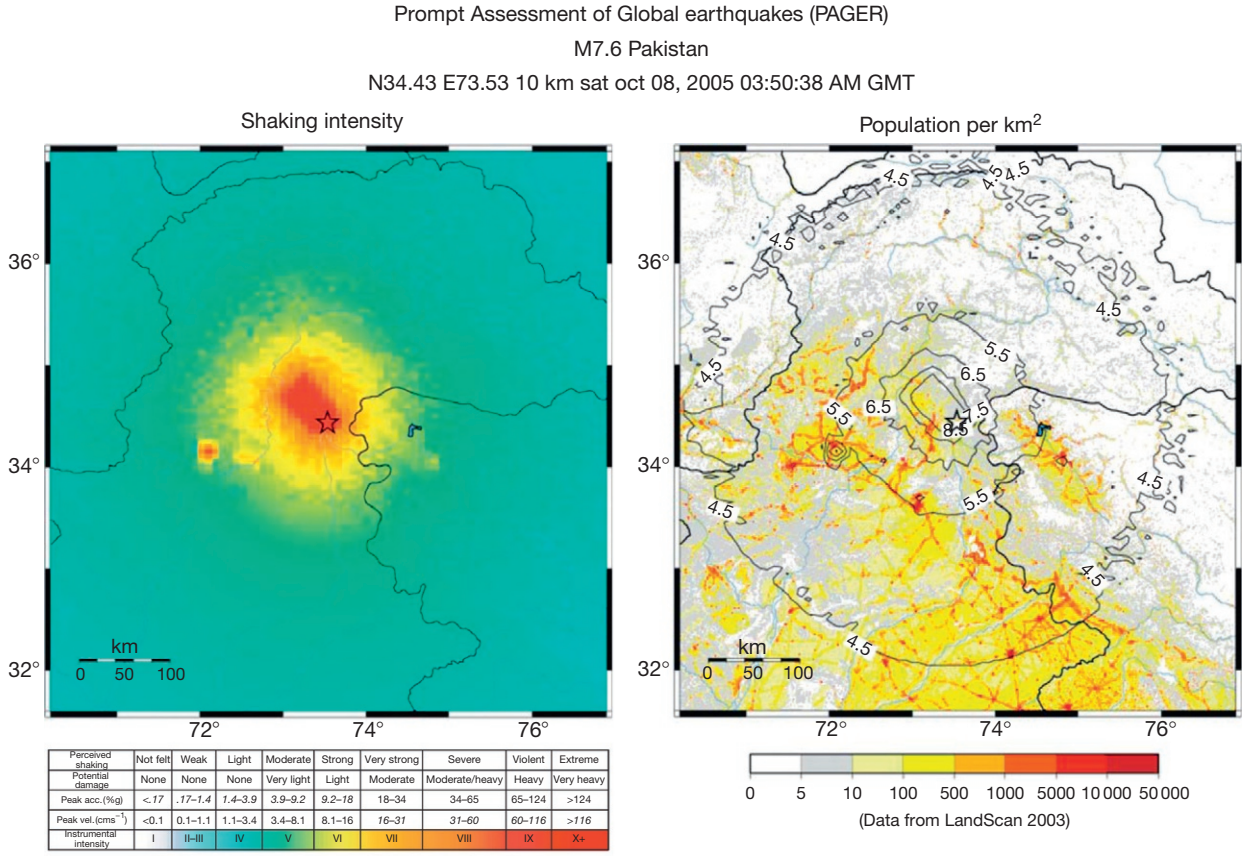


Perceived shaking	Not felt	Weak	Light	Moderate	Strong	Very strong	Severe	Violent	Extreme
Potential damage	None	None	None	Very light	Light	Moderate	Moderate/heavy	Heavy	Very heavy
Peak acc. (%g)	*<.17*	*.17–1.4*	*1.4–3.9*	*3.9–9.2*	*9.2–18*	18–34	34–65	65–124	>124
Peak vel. (cm s^{-1})	<0.1	0.1–1.1	1.1–3.4	3.4–8.1	8.1–16	*16–31*	*31–60*	*60–116*	*>116*
Instrumental intensity	I	II–III	IV	V	VI	VII	VIII	XI	X+

Figure 12 Scenario ShakeMap for an *M* 7.3 rupture of the Hayward–Rodgers Creek Fault. This is one of the earthquake rupture scenarios identified by WG02 (2003) and was assigned a 1% probability of occurrence by 2032.

seismic 'risk,' that is, the impact of an earthquake, which is more desirable for most consumers. For emergency services personnel, they respond to locations where the greatest hazard intersects the built environment. ShakeCast is intended to provide sophisticated users with the necessary tools to assess the most likely damage to facilities provided the fragility is known. In an ambitious new project, the USGS National Earthquake Information Center (NEIC) is developing a methodology to convert ground shaking hazard into an assessment of impact on the local population. The Prompt Assessment of Global Urban Earthquakes for Response (PAGER) methodology aims to first estimate the distribution of ground shaking and then estimate the number of fatalities (Earle et al., 2005).

To estimate the distribution of ground shaking, that is, a ShakeMap, for a global event, the minimum required data are the earthquake location and magnitude, which are routinely determined for global earthquakes with $M > 5$ by the NEIC. Using available attenuation relations and site corrections derived from the local topography, an initial estimate of the distribution of ground shaking can be made. Additional data that can be input as available include recorded local ground motions, ground shaking intensities reported through the CIIM system, and information about fault finiteness. The finite-source information can be derived from a range of sources including aftershock distributions, broadband waveform inversion of teleseismic data (e.g., Ji et al., 2004), and field observation in the hours and days after an event (Wald et al., 2006a). Combining the ShakeMap with population distribution, the number of people experiencing ground shaking at various intensities can be estimated. **Figure 13** shows an example of the PAGER output for the 2005 M_W 7.6 Pakistan earthquake. The methodology estimates that almost 10 million



Population exposed to shaking

MMI instrumental intensity population

IX	587 000
VIII	1 200 000
VII	1 860 000
VI	9 460 000
V	37 300 000*

Figure 13 Output from the USGS National Earthquake Information Center's prototype PAGER system. ShakeMap for the event is shown at upper left, and MMI contours are overlain on the population density (upper right). The number of people experiencing different levels of ground shaking can then be tabulated, below. More information is available at http://earthquake.usgs.gov/eqcenter/pager/.

people experienced an MMI of VI and 587 000 experienced an MMI of IX. Ongoing development of PAGER aims to provide regional fragility information so that these figures can be converted into estimates of the number of casualties.

4.23.5.2 Warnings Before the Shaking

The tools and methodologies described earlier in the text provide rapid postearthquake information in the minutes to hours after an event. This information is critical to the emergency response community and can prevent cascading failures. It is also useful for longer-term planning and training purposes. But the rapid earthquake information system first described by Cooper (1868) envisioned a warning system designed to provide an alarm prior to ground shaking. Such warning systems could be used for short-term mitigation in the seconds to tens of seconds prior to ground shaking to prevent damage, casualties, and fatalities. The scientific and engineering challenge for any such warning system is to rapidly distinguish between the frequently occurring small and harmless earthquakes and the large damaging ones.

4.23.5.2.1 S-waves versus P-waves

The simplest warning system monitors ground motion and issues an alert or mitigating action when the ground acceleration exceeds some critical threshold. The thresholds are set high, typically $\sim 0.04g$ (where g is the acceleration due to gravity), which is the level at which buildings and other infrastructure start to experience permanent damage. These systems therefore not only trigger on S-wave energy and have a zero warning time but also have the benefit that there is no prediction required; the critical ground shaking has been observed when the alert is issued. Such ground shaking detectors are used widely to shut down utility, transportation, and manufacturing systems during earthquakes.

These detectors can be turned into a true warning system, that is, greater than zero seconds warning, by placing them between the earthquake source and the infrastructure or city they are intended to protect. The warning is then transmitted ahead of ground motion electronically. This 'front-detection' approach is being used in Japan and Mexico, where subduction zone earthquakes along the Japan and Middle America Trenches represent a significant hazard for cities further inland. By deploying stations along the coastline adjacent to the earthquake source region, warning can be transmitted electronically ahead of the ground shaking (Espinosa-Aranda et al., 1995; Nakamura and Tucker, 1988). A nonzero warning time requires some form of prediction as ground motion parameters must be detected at one location and estimated for another; this introduces uncertainty. In the case of front detection, ground motion parameters close to the epicenter are used to predict ground shaking levels further away. When the geography is conducive, these systems can provide substantial warning times. In the case of the Seismic Alert System (SAS) in Mexico, the ~300 km between the subduction zone and Mexico City provides for ~70 s of warning as was demonstrated in the 1995 M_W 7.4 Guerrero earthquake (Anderson et al., 1995).

The amount of warning can be increased by using the P-wave rather than the S-wave energy to assess the magnitude or hazard associated with an earthquake. Nakamura (1988) first proposed such an approach that was implemented along the Shinkansen (bullet train) lines in Japan in the 1990s. Nakamura's approach is to use the predominant period, that is, the frequency content, of the first few seconds of the P-wave to estimate the magnitude of an earthquake. For seismic stations within ~150 km, this measurement is relatively insensitive to epicentral distance and geographic location. Observations from the first few seconds of P-waves recorded within ~150 km of the epicenter of $3 \leq M \leq 8.3$ earthquakes around the world show a scaling relation between magnitude and frequency content, τ_p^{max}, as shown in **Figure 14** (Olson and Allen, 2005). This provides one basis for an early-warning system. The hazard posed by an earthquake is expressed in terms of the magnitude estimate derived from τ_p^{max} of P-waves recorded close to the epicenter. There is uncertainty in magnitude estimates derived from this relation. In the case of the global data set (**Figure 14**), it is ±1 magnitude unit, although these uncertainties can be reduced as discussed later in the text. Similar magnitude–frequency scaling relations have been developed for various regions around the world (Allen and Kanamori, 2003; Kanamori, 2005; Lockman and Allen, 2007; Nakamura, 2004; Simons et al., 2007; Wu and Kanamori, 2005a,b), although the approach also has its detractors (e.g., Rydelek and Horiuchi, 2006; see also Olson and Allen, 2006 response).

In addition to using the frequency content of the P-wave, the amplitude can also be used to assess the forthcoming hazard associated with the S- and surface-wave energy. Wu and Kanamori (2005a,b) explored the use of the peak displacement, velocity, and acceleration within the first 3 s of the P-wave. They found that the lower-frequency content of the peak displacement has a high correlation with the peak ground displacement (PGD) and the PGV observed many seconds later. **Figure 15** shows the relation between P_d, the PGD observed within 3 s of

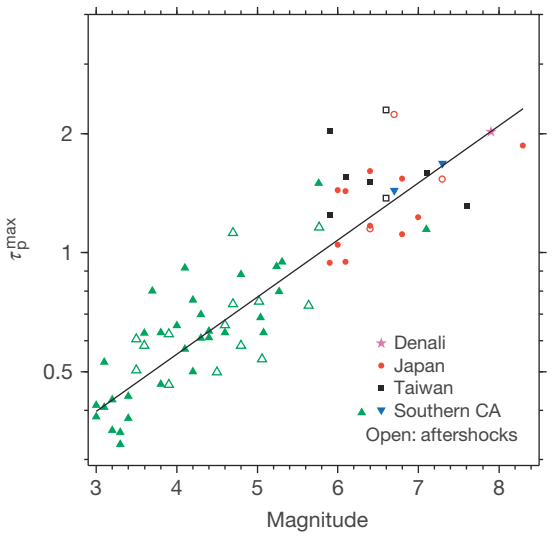


Figure 14 Scaling relation between earthquake magnitude and the frequency content of the first 4 s of the P-wave recorded at stations within 150 km. This global data set consists of 1842 waveforms recorded from 71 earthquakes. The individual values of τ_p^{max} at each station are averaged on this plot. All the event-averaged values fall within a range of ±1 magnitude unit. Modified from Olson E and Allen RM (2005) The deterministic nature of earthquake rupture. *Nature* 438: 212–215.

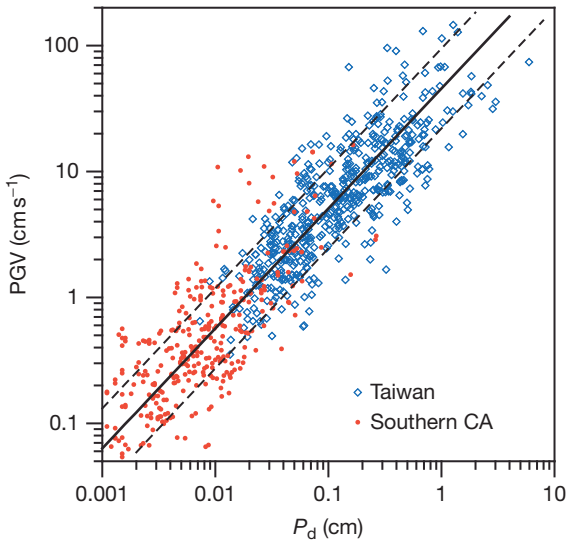


Figure 15 Scaling relation between P_d (the peak displacement observed within the first 3 s of the P-wave) and PGV observed at the same station. Data from 38 $M \geq 5.0$ earthquakes in Taiwan (blue) and southern California (red) are shown. Modified from Wu Y-M, Kanamori H, Allen RM, and Hauksson E (2007) Experiment using the tau-c and Pd method for earthquake early warning in southern California. *Geophysical Journal International* 170: 711–717.

the P-wave arrival, and PGV for 38 $M \geq 5.0$ earthquakes from Taiwan and southern California (Wu et al., 2007). P_d observations at a site can therefore be used to assess the forthcoming ground shaking hazard at the same site. P_d, and similar amplitude-derived parameters, can also be used to estimate earthquake magnitude once corrected for attenuation associated with the epicentral distance (Kamigaichi, 2004; Odaka et al., 2003; Wu and Kanamori, 2005b; Wu and Zhao, 2006; Wurman et al., 2007). In a novel hybrid approach, Cua (2005) used the amplitude of waveform envelopes to estimate the magnitude of an earthquake. The magnitude determination is derived from the ratio of the peak P-wave displacement and acceleration. Given the different frequency sensitivities of the acceleration and displacement waveforms, this approach is analogous to the predominant period approach first suggested by Nakamura, but was arrived at independently using a linear discriminate analysis.

4.23.5.2.2 Single-station and network-based warnings

The simplest and most rapid approach to providing a ground shaking warning is to use a single seismic station to record ground motion parameters and issue a warning on site. The UrEDAS first outlined by Nakamura (1988) provides an estimate of the magnitude and location of an earthquake using just a single three-component seismometer. Criteria for taking mitigating actions are then developed based on the expected peak ground shaking and warning time, which are derived from the magnitude and epicentral distance of the event. Alternatively, rather than first estimating the magnitude, the hazard at the station site can be estimated directly. **Figure 15** is an example of this where PGV is estimated directly from P_d. Combining the amplitude and frequency information from P-waves for $M \geq 5.0$ earthquakes in Taiwan, Wu and Kanamori (2005b) showed that the sites that later experienced damaging ground motion could be distinguished from those that did not. The advantage of this approach is its speed. With this approach, it is possible to provide warning at the epicenter. As soon as information about an earthquake is available at a site, action can be taken. The disadvantage, compared with a multiple-station approach, is greater uncertainties in the hazard estimates and the warning time; in some cases, no estimate of the warning time is available. However, the choice of appropriate sites for single-station systems can significantly improve their accuracy. Lockman and Allen (2005) applied a similar methodology to UrEDAS to all broadband velocity stations in southern California. They found one quarter of the stations produced magnitude estimates with errors less than ± 0.3 magnitude units, hypocentral distances within ± 15 km, and back azimuth calculations within $\pm 20°$, but the errors at other stations were larger, making some unusable for the purpose of early warning.

A network or regional-based approach is the alternative to single-station systems. By combining information from multiple stations, the uncertainties in hazard estimates and the number of false alarms can be reduced. Network-based approaches typically locate an earthquake and estimate its magnitude as a first step to predicting the expected distribution of ground shaking (Allen, 2007; Allen and Kanamori, 2003; Cua, 2005; Horiuchi et al., 2005; Kamigaichi, 2004; Wu and Teng, 2002; Wurman et al., 2007). The site-specific peak ground shaking and the time at which it is expected can then be transmitted to users to initiate mitigating actions. When compared with a single-station approach, the cost for users close to the epicenter is a reduced warning time as the system must wait for seismic arrivals at multiple seismic stations and data must be telemetered between sites. However, the introduction of a regional telemetry system increases warning times for users further from the epicenter. For an earthquake detected close to the epicenter, the warning can be transmitted ahead of the ground shaking. This is the front-detection approach described in the preceding text.

4.23.5.2.3 Warning around the world

It is clear that the most accurate and timely, that is, the most effective, warning systems will combine all of the previously mentioned approaches making use of information contained in the full waveform and issuing warnings on site and taking advantage of a network and telemetry system. **Figure 16** shows the locations of the warning systems now in operation and development around the world. Most make use of hybrid methodologies.

The operational systems are in Japan, Taiwan, Mexico, and Turkey, where warnings are issued to users beyond the seismological community. In Japan, the first alarm seismometers were deployed by Japan Railways Group in the mid-1960s (Nakamura and Tucker, 1988); these detectors were then developed into the more sophisticated UrEDAS P-wave detection system (Nakamura, 1988) in the early 1990s. Since then, network-based approaches have been developed by both the JMA (Kamigaichi, 2004) and the NIED (Horiuchi et al., 2005). JMA has been testing an early-warning system for general use since February 2004 (Kamigaichi, 2004). In August 2006, they widened the testing to 41 institutions, including railway companies, construction firms, factories, and hospitals. As the

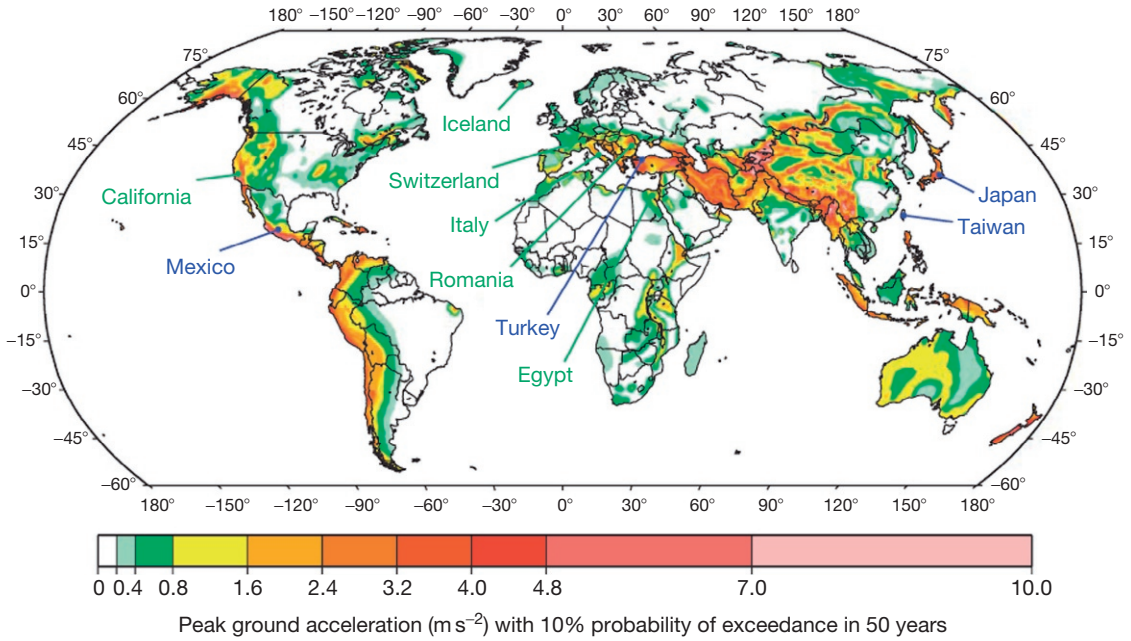


Figure 16 Map showing the locations of earthquake early-warning systems currently in operation (blue) or development (green) around the world. Operational systems include Japan, Taiwan, Mexico, and Turkey. Systems are in development for California, Egypt, Greece, Iceland, Italy, Romania, and Switzerland. An operational system is defined as one that issues warning information to users outside the seismological community. The locations are overlaid on the GSHAP map. Reproduced from Giardini D, Grunthal G, Shedlock KM, and Zhang PZ (1999) The GSHAP global seismic hazard map. *Annali Di Geofisica* 42: 1225–1230.

public becomes more familiar with the system, they plan to make the information more widely available.

The Central Weather Bureau in Taiwan has been using a virtual subnet approach to rapidly assess magnitude from the S-wave energy of an event. This method requires an average of 22 s for magnitude determination and gives warning to populations greater than 75 km away (Wu et al., 1998; Wu and Teng, 2002). The development by Wu and colleagues of P-wave methodologies described earlier in the text is aimed at increasing the warning times and reducing the blind zone where warnings cannot be provided. Using a network approach, it is estimated that the blind zone would be reduced to 20 km (Wu and Kanamori, 2005a). Single-station methodologies could provide warnings at smaller epicentral distances (Wu et al., 2006).

Mexico City's SAS takes advantage of its geographic separation from the seismic source region along Guerrero gap subduction zone to the southwest. The front-detection system measures the rate of increase of S-wave energy at stations along the coast to estimate magnitude and transmits this information to the population in Mexico City 300 km away (Espinosa-Aranda et al., 1995). It has been operational since 1991 and transmits its warnings to schools, industry, transportation systems, and government agencies. Finally, Turkey is the most recent member of the early-warning club. Their system triggers when the amplitude of ground motion exceeds some threshold at a network of instruments around the Sea of Marmara, providing warning to users in Istanbul (Boese et al., 2004; Erdik et al., 2003).

The development of early-warning systems is also underway across Europe and in the United States. The European Community is currently funding the cooperative development and testing of early-warning algorithms in Egypt, Greece, Iceland, Italy, Romania, and Switzerland. In the United States, the California Integrated Seismic Network (CISN) has recently embarked on a project to test various early-warning algorithms to evaluate their performance across the state. The test includes two network-based approaches, the Earthquake Alarm Systems (ElarmS) and the Virtual Seismologist (Cua, 2005), and a single-station approach, the amplitude, and period monitor (Wu and Kanamori, 2005a). The goal is to evaluate the real-time performance and strengths of these methodologies in order to develop an optimal hybrid system for the state. In order to get a sense of the capabilities of such a future system, we consider the performance of one of these methodologies, the one most familiar to the author, in more detail.

4.23.5.2.4 ElarmS in California

The Earthquake Alarm Systems (ElarmS) is a network-based approach to earthquake early warning (Allen, 2004, 2007; Allen and Kanamori, 2003; Wurman et al., 2007; http://www.ElarmS.org). The methodology uses the first 4 s of the P-wave arrival at stations in the epicentral region to locate earthquakes in progress and estimate their magnitude. An AlertMap is generated, showing the expected distribution of peak ground shaking in terms of PGA, PGV, and MMI. All available data are collected from all stations every second and the AlertMap is updated. Initially, the AlertMap is based on the location and magnitude estimates only, and an attenuation relation is used to predict ground shaking. As time proceeds, observations of peak ground shaking near the epicenter are incorporated into the estimate of ground shaking at more distant locations. The predictive AlertMap therefore evolves into an observed ShakeMap during the course of an event.

The ElarmS algorithms were developed using calibration data sets for both southern California and northern California. Since February 2006, they have been automatically processing all $M \geq 3.0$ earthquakes in northern California. They are not yet part of the real-time system and are running in an off-line mode. On notification of an earthquake from CISN, they sleep for 10 min to allow waveform data to populate the archive. They then gather all available data and process these data without human interaction to generate a time series of AlertMaps. Between February 2006 and September 2006, there were 83 events processed in this fashion. Figure 17 shows the AlertMap output for one of the largest events during this period, the M_L 4.7 earthquake near Santa Rosa on 2 August 2006 (local time). The time histories of the magnitude, PGA, PGV, and MMI prediction errors are shown in Figure 18. This event was near the Rodgers Creek Fault in a similar location to one of the future hazardous scenario events in the region (WG02, 2003).

The initial detection occurs 3 s after the event origin time (Figure 17(a)). The event is located (red star) at the station to trigger (gray triangle) and the warning time across the region is estimated (concentric circles). One second later (Figure 17(b)), an additional two stations trigger and the event is relocated using the grid search method. The initial magnitude estimate is also available, derived from the first second of data from the first station to trigger.

The initial estimate is high, M 5.8, and the predicted distribution of peak ground shaking is correspondingly high (color pallet). The MMI estimates exceed the actual observations by up to 2 MMI units. One second later (Figure 17(c)), magnitude estimates are available from the additional two triggered stations providing an updated event magnitude estimate of M 4.3. This reduces the predicted MMI intensities and reduces the errors in all output parameters (Figure 18). This illustrates the benefit of using multiple stations. In this case, waiting one additional second so that magnitude information is available from three stations rather than one station significantly reduces the error.

One second later, just 3 s after the initial detection, peak ground shaking is observed at two stations (Figure 17(d) – colored triangle and diamond), and these observations are used to adjust the attenuation relations for the region. While

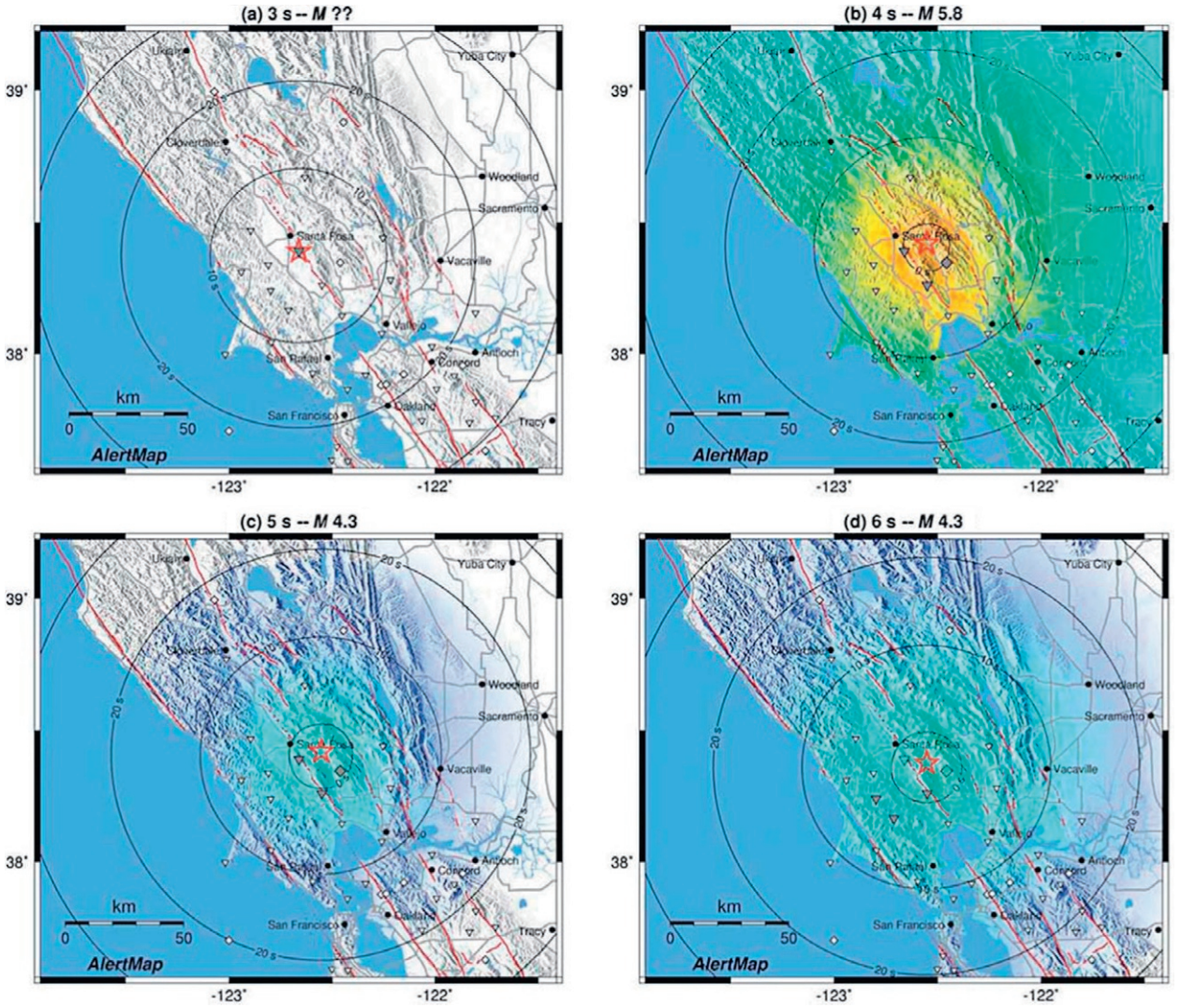
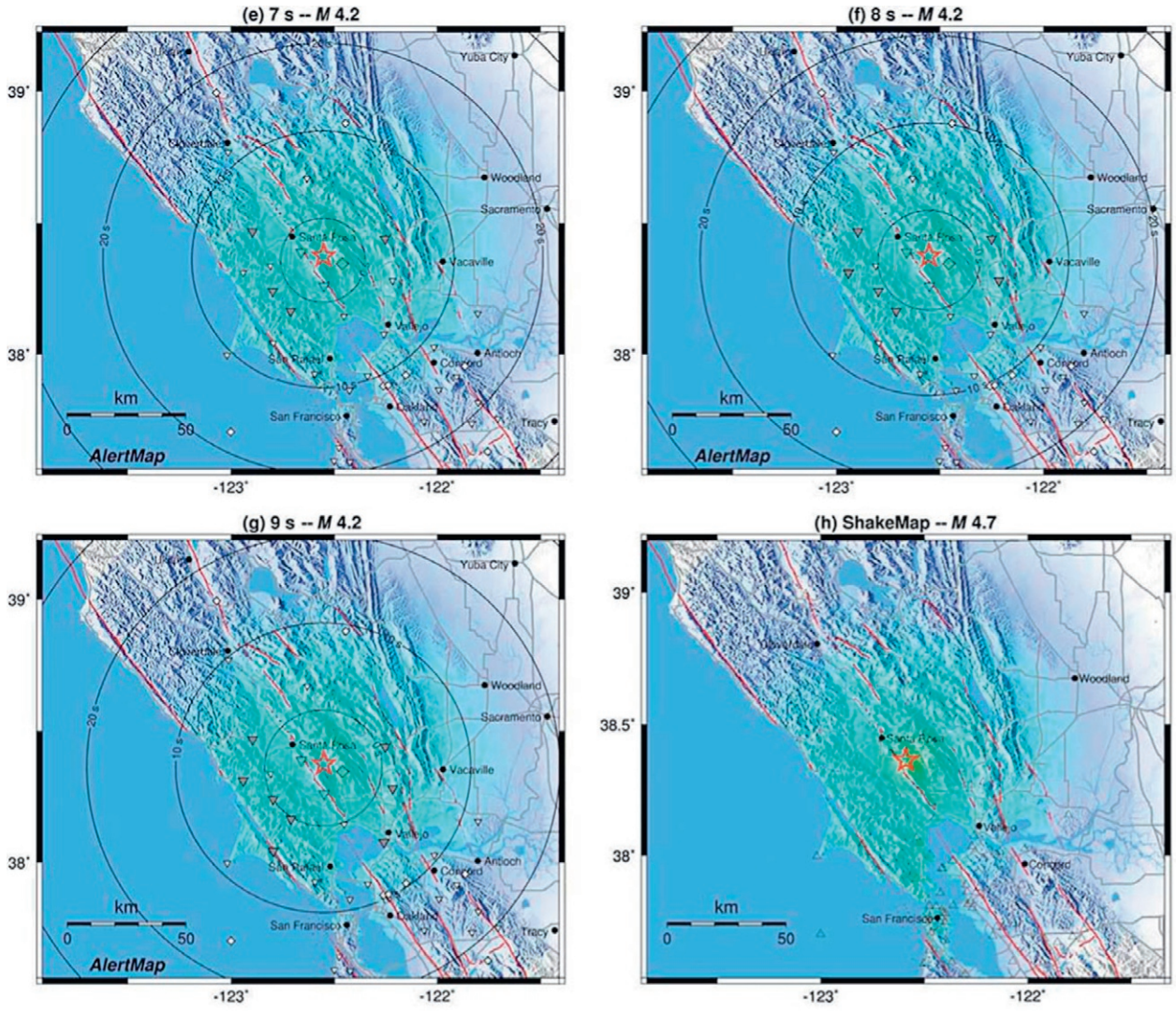


Figure 17 (Continued)



Perceived shaking	Not felt	Weak	Light	Moderate	Strong	Very strong	Severe	Violent	Extreme
Potential damage	none	none	none	Very light	Light	Moderate	Moderate/Heavy	Heavy	Very Heavy
Peak acc. (%g)	*<.17*	*.17-1.4*	*1.4-3.9*	*3.9-9.2*	*9.2-18*	18-34	34-65	65-124	>124
Peak vel. (cm/s)	<0.1	0.1-1.1	1.1-3.4	3.4-8.1	*8.1-16*	*16-31*	*31-60*	*60-116*	*>116*
Instrumental intensity	I	II-III	IV	V	VI	VII	VIII	IX	X+

Figure 17 Performance of ElarmS for the M_L 4.7 earthquake near Santa Rosa on 2 August 2006 (local time). (a–g) AlertMap output from the time of initial detection, 3 s after event origin time, for 8 consecutive seconds. (h) The event ShakeMap for comparison. The red star is the event epicenter; concentric circles indicate the warning time. Triangles (broadband velocity), inverted triangles (strong motion), and diamonds (collocated velocity and strong motion) show the locations of seismic station. The symbols turn gray when the station triggers and are colored according to the peak ground shaking at the site once it has occurred. The color pallet shows the predicted instrumental MMI for the AlertMaps (a–g) and the 'observed' for the ShakeMap (h).

the magnitude estimate remains 0.4–0.5 units low for the following 6 s (Figure 18(a)), the effect on ground shaking estimates is reduced by the inclusion of these peak ground shaking estimates at the closest stations (Figure 18(b)–18(d)). AlertMaps for the following 3 s are shown in Figure 17(e)–17(g). Additional stations trigger providing information for the magnitude estimate, and peak ground shaking is observed at additional sites, but the predicted distribution of ground shaking does not change noticeably. The CISN ShakeMap for this event is shown in Figure 17(h) for comparison. The AlertMap from 6 s onward is very similar, the main difference being the slightly stronger ground shaking at the epicenter on the

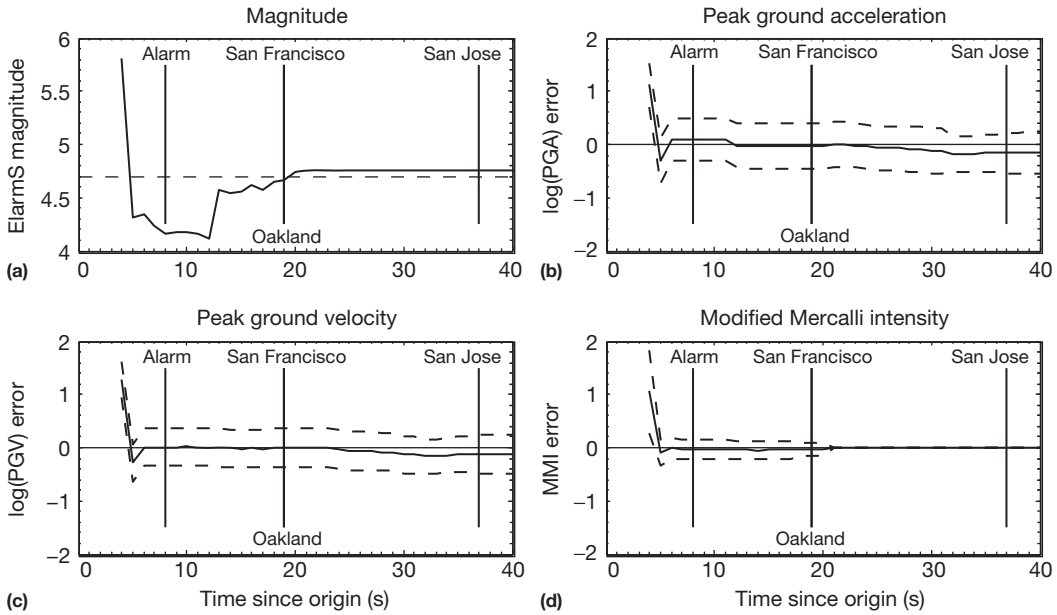


Figure 18 Performance of ElarmS for the M_L 4.7 earthquake near Santa Rosa on 2 August 2006 (local time) as a function of time. (a) ElarmS magnitude estimate; the dashed line is the California Integrated Seismic Network magnitude of M_L 4.7. (b) Errors in the predicted PGA determined by subtracting the logarithm of the observed from the logarithm of the predicted. Only stations where the peak ground shaking has not yet been observed are included. The dashed lines represent the one-sigma error envelope. (c) Errors in PGV. (d) Errors in MMI. The MMI error goes to zero, as all stations that have not yet observed peak ground shaking after 20 s had a predicted MMI value of *I* and an observed value of *I*. The vertical bars indicate the alarm time (4 s of P-wave data available from four sensors) and the time of peak ground shaking in the cities of San Francisco, Oakland, and San Jose.

ShakeMap. This is due to the underestimates of the ElarmS magnitude, which remains low until 13 s, when it reaches *M* 4.6. Details of the ElarmS methodology and performance in northern California can be found in Wurman et al. (2007).

The continuum of information available about an ongoing earthquake is illustrated in **Figure 18**, which shows the changing error in the predictions. Any individual user can decide whether they would rather react to earlier information that has not only greater uncertainty but also greater warning time or wait a few seconds for the uncertainty to reduce. This decision can be made in a probabilistic framework (Grasso, 2005; Grasso and Allen, in review; Iervolino et al., 2006). When the cost of inaction in a damaging earthquake and the cost of taking mitigating action are known, the appropriate predicted ground shaking threshold for talking action can be defined provided the uncertainty in the prediction is also known. By only taking action when this threshold is reached, the total cost of an earthquake is minimized.

4.23.5.2.5 Warning times

The maximum warning time for the Santa Rosa event is 15 s for San Francisco and Oakland and 33 s for San Jose in the South Bay. This is the time from the initial magnitude estimation until maximum ground shaking in the cities. However, the initial prediction is high, so it would be preferable to wait at least a few seconds before taking any actions. The 'alarm time' is defined in this chapter as the time at which 4 s of P-wave data is available from four seismic instruments. Application of ElarmS to data sets from southern California, northern California, and Japan shows that the average absolute magnitude error at this time is 0.5 units (Allen, 2007; Wurman et al., 2007). The alarm time for the Santa Rosa event is shown in **Figure 18**; from alarm time, there are still 11 s warning for Oakland and San Francisco and 24 s for San Jose. A second M_L 4.7 earthquake occurred in northern California since the automated ElarmS processing began. It occurred on 15 June 2006 near Gilroy, south of the bay, and was almost the same distance from San Francisco and Oakland as the 1989 Loma Prieta earthquake. At alarm time for the Gilroy event, when the magnitude estimate was 4.3, there was 3 s of warning for San Jose, 20 s warning for Oakland, and 22 s for San Francisco. In the Loma Prieta earthquake, 84% of the fatalities occurred in Oakland and San Francisco. Therefore, in a repeat of the Loma Prieta earthquake with a warning system in place, there could be ~20 s warning in the locations where most casualties occur.

Warning times for earthquakes in California range from zero seconds up to over a minute depending on the location of the earthquake with respect to a population center. Heaton (1985) used a theoretical distribution of earthquakes in southern California to estimate the range of warning times as a function of ground shaking intensities at the warning location. He showed that for the larger, most damaging earthquakes, there could be more than 1 min of warning.

Using the ElarmS methodology, we can estimate the warning time for any earthquake location. **Figure 6** contours the warning time the city of San Francisco would have for an earthquake with an epicenter at any location across the region. The warning time is the difference between the alarm time for the earthquake given the current distribution of real-time seismic stations and the time at which peak ground shaking would

occur in San Francisco. An additional 5.5 s has been deducted from the warning time to account for telemetry delays of the existing network (which could be reduced).

While Figure 6 shows the warning time for all earthquake locations, future damaging events will likely occur on specific faults. These likely future damaging earthquake scenarios were identified by WG02. As probabilities are associated with each earthquake scenario, probabilities that an earthquake with a particular warning time will occur by 2032 can be estimated. Figure 19 shows that distribution of the warning times for these scenario earthquakes ranges from −7 to 77 s where a negative warning time means that the alert time was after the peak ground shaking in San Francisco. The most likely warning times range from −7 to 25 s, which are due to earthquakes on the numerous faults throughout the SFBA (Figure 6). The long tail extending to 77 s is due to events on the San Andreas extending to the north. The scenario ShakeMaps for each event (e.g., Figure 13) provide an estimate of the ground shaking intensity in San Francisco. The probability distribution shown in Figure 19 is colored accordingly. The inset in Figure 19 shows the probability that there will be more or less than 0, 5, 10, 20, and 30 s of warning and shows that it is more likely that there will be more than 10 s of warning for the most damaging events. If the telemetry delay was reduced, or more stations were deployed to the north of the SFBA, then more than 20 s warning is likely for these most damaging earthquakes. One of the most deadly scenarios for the city of San Francisco is an *M* 8, 1906-type earthquake, with a rupture initiating near Cape Mendocino and propagating south. In this scenario, there could be over 1 min of warning time. Probabilistic warning time distributions for various other locations are also available (Allen, 2006).

4.23.5.2.6 Future development

The large-magnitude, most damaging earthquakes are when a warning is of most value and also when the warning times can be the greatest. The accuracy of the ground shaking predictions for these large-magnitude events is significantly improved by knowledge of the finiteness of the rupture. Neither ElarmS nor any of the other operational early-warning systems currently account for fault finiteness. This is therefore an active area of research. One approach is to monitor the displacement across fault traces allowing instantaneous identification of rupture. This requires instrumentation along all faults and also that the rupture occurs on a previously identified fault at the surface. Some of the earliest proposals for warning systems used wires across fault traces to detect slip. Today, real-time GPS stations could be used to monitor displacement and would be sensitive

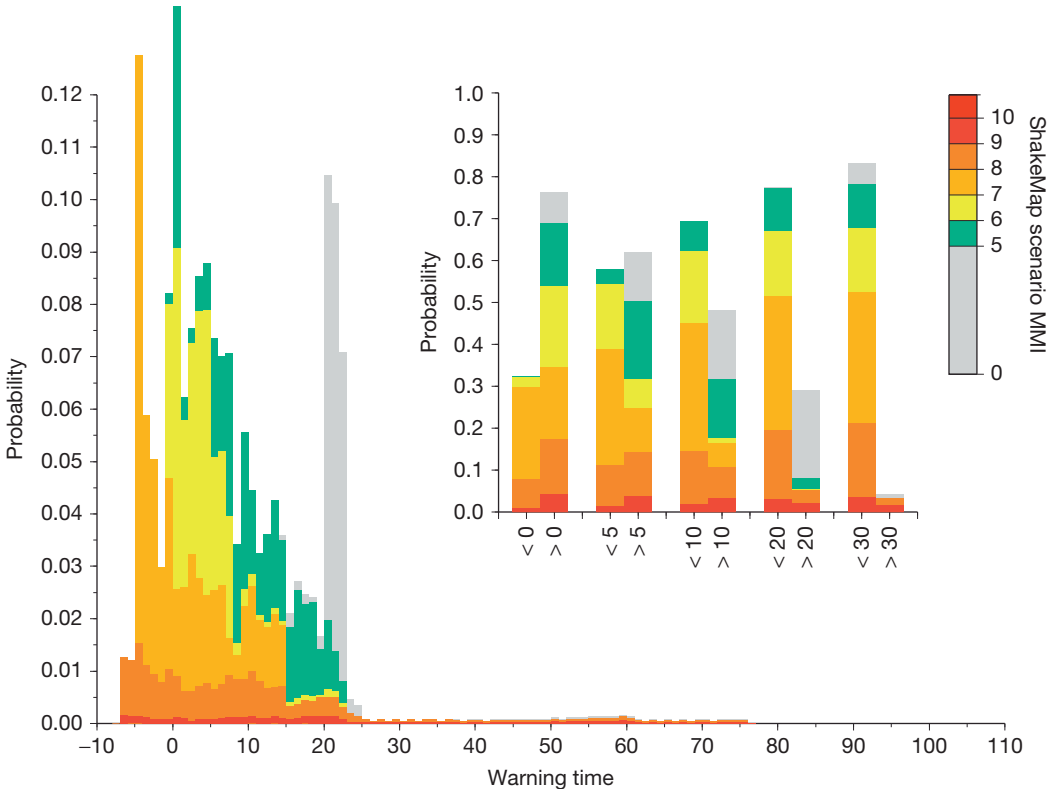


Figure 19 Warning time probability density function for the city of San Francisco. The warning times for all earthquake scenarios identified by WG02 were estimated given the current seismic network and telemetry delays using ElarmS. The range of warning times is −7 to 77 s where a negative warning time means that peak ground shaking occurs before the warning is available. The most probable warning times range from −7 to 25 s; the long tail extending to 77 s is due to the San Andreas Fault. The color shows the predicted intensity of ground shaking in the city. The inset shows the probability of more or less than 0, 5, 10, 20, and 30 s warning. It is much more likely there will be greater than zero seconds warning, and the warning times are greater for the most damaging earthquakes. Modified from Allen RM (2006) Probabilistic warning times for earthquake ground shaking in the San Francisco Bay Area. *Seismological Research Letters* 77: 371–376.

to slip on fault planes at greater distances. An alternative approach is identifying which seismometers are near-field and which are far-field during the rupture in order to map the rupture extent. Yamada and Heaton (2006) used the radiated high-frequency energy at near-field stations to approximate the rupture area and the evolving moment magnitude in order to estimate the probable rupture length. As these real-time finite-fault techniques are developed, it will be important to incorporate them into early-warning systems.

4.23.5.2.7 Benefits and costs

Warning information from the operational warning systems in Japan, Taiwan, Mexico, and Turkey are currently used by transportation systems such as rail and metro, as well as private industries, including construction, manufacturing, and chemical plants. They are also used by utility companies to shut down generation plants and dams and emergency response personnel to initiate action before ground shaking. In addition, schools receive the warnings allowing children to take cover beneath desks, housing units automatically switch off gas and open doors and windows, and entire complexes evacuate. These same applications would be appropriate for early-warning implementations in many regions around the world and include both automated response by a computerized control system and human response (for both personal protection and reduction of damage to infrastructure).

Looking to the future, earthquake engineering is already evolving to incorporate real-time earthquake information from early-warning systems. In Japan, most new high-rise buildings are 'dynamic intelligent buildings' that contain structural control devices to select or change the vibration characteristics of a building, that is, the stiffness or damping (e.g., Housner et al., 1997). Some of these buildings have active control systems that use external power to change or control the building's response to vibrations. Others have passive devices that use hysteretic or viscoelastic properties of material to reduce vibrations with no external power. More recently, semiactive systems have been developed that use passive devices that are actively put into operation when necessary. Early-warning information is of value to both the active and semiactive types. As more information about the characteristics of forthcoming ground shaking becomes available (such as amplitude and frequency content), the more effective the building's response systems can be.

For personal protection, early-warning systems could perhaps be of most value in regions with high seismic hazard and poor implementation of earthquake-resistant building practices. In many of these underdeveloped environments, buildings are typically small single-story dwellings. Homes may be built by the owner using local materials such as mud bricks. Earthquakes in these regions have high-fatality rates as buildings collapse on their occupants. For example, the recent 2003 Bam (Iran) and 2005 Pakistan earthquakes together killed over 100 000 people. In these environments, it only takes a few seconds to get out of these buildings, and early-warning systems could provide that time.

The costs of early-warning systems are substantial, but so are the costs of other mitigation strategies and the earthquakes themselves. California currently has ~300 seismic stations that are telemetered in real time and appropriate for use in an early-warning system. Broad implementation of earthquake early warning in the region would require a more robust and redundant seismic network. To install an additional 600 instruments would cost between $6 million and $30 million, depending on the instrumentation used. To operate that network would cost between $2 million and $6 million per year. In addition to these costs, a system to transmit the warning information would be needed as well as an educational program to teach people how to use the information. For comparison, UC Berkeley is currently retrofitting campus buildings to prevent collapse in future earthquakes. The cost per building is typically $10–$30 million; retrofit of the historical Hearst Memorial Mining Building costs $80 million and was made possible by a generous donation. UC Berkeley is spending $20 million per year for 20 years to protect its students and staff in an earthquake, and indeed its very own existence, against a significant earthquake in the region. Implementation of an early-warning system in California is not a replacement for earthquake-resistant buildings and retrofit programs, but there are hundreds of buildings in the SFBA alone like those currently being retrofit on the Berkeley campus that will not be retrofitted. An early-warning system would allow some short-term mitigation strategies for everyone.

Similarly, in regions where there is little or no implementation of earthquake-resistant building practices, a warning system would provide some mitigation of earthquake effects. The costs could perhaps be reduced by using clusters of stations to improve on single-station performance without requiring a full seismic network. The operation of such systems would have to be done locally, requiring a local seismological skill base. Developing this skill base will also perhaps assist in the improvement of building practices, so both long-term building and short-term warning can be used to reduce the costs of future earthquakes.

4.23.6 Conclusion

Progress in seismic hazard mitigation has been substantial – near-zero fatalities from all earthquakes are within our technical capabilities – and yet the cost of earthquakes is still rising, and the number of fatalities continues to increase.

Reducing the cost and fatalities in future earthquakes requires first identifying the hazard and then implementing appropriate mitigation strategies. Our understanding of the earthquake process allows effective long-term forecasts of hazard expressed as the probability of ground shaking above some threshold. Plate tectonics provides the framework for understanding where most future earthquakes will occur. When considered as a stationary time series, the likelihood of future events can be estimated with a degree of confidence. This provides earthquake probability forecasts on timescales of fifty to hundreds of years. Yet most in the seismological community would agree that there is a time dependence to earthquake hazard and the probability of a large earthquake increases with time since the last event as stress increases on a fault. The challenge is to estimate the likely time until the next rupture, which is dependent on not only the rate of increasing stress but also the initial stress, activity on surrounding faults, and changes in the physical properties of the crust. Given these

limitations, the uncertainty in hazard forecasts increases as the forecast timescale decreases.

While the public continues to identify short-term earthquake prediction – the high probability of a clearly defined earthquake in a short period of time – as the solution to earthquake disasters, few seismologists see such predictions as feasible within the foreseeable future. Existing mitigation strategies, when fully implemented, could reduce the impact of earthquakes more than even the most accurate short-term predictions. This is because predictions would only allow people to get out of the danger area, but the infrastructure on which their lives depend would remain.

Mitigation strategies fall into two categories: long-term and short-term. Long-term mitigation focuses on building infrastructure capable of withstanding earthquake shaking. This approach has been very effective in reducing the number of fatalities in earthquakes, but still new lessons are learned each time there is a large damaging earthquake. New techniques now allow engineers to test designs against the shaking anticipated from future earthquakes. This provides the opportunity to move beyond the current mode dominated by response to what did not work in the last earthquake. PBSD is now also providing a framework for reducing the economic impacts of earthquakes in addition to preventing fatalities.

Short-term mitigation is provided by rapid earthquake information systems. Modern seismic networks have been providing location, magnitude, and ground shaking information in the minutes after an event for over a decade. This information has now been widely integrated into emergency response, allowing for more efficient and effective rescue and recovery efforts. But today, many earthquake-prone regions are pushing the limits of rapid earthquake information systems in an effort to provide similar information in the seconds to tens of seconds before the ground shaking. These warning systems provide another opportunity to further reduce the costs and casualties in future earthquakes.

The reduction of seismic risk will be most effective when multiple approaches are used. There is still a surprise component to all large-magnitude earthquakes, which acts as a reminder that we need to be wary of becoming too tuned in our mitigation efforts. By combining earthquake-resistant design to prevent building collapse, warning systems to isolate toxic systems, and rapid response to critical facilities identified as potentially damaged, we can reduce the impact of an earthquake and also accommodate the failure of one component in the system. In another situation, one mitigation strategy might not be economically feasible while another is. It is therefore important to continue development of a full range of methodologies.

Perhaps the greatest challenge in seismic hazard mitigation is the implementation of these mitigation strategies in all earthquake-prone regions. While the hazard is now clearly identified on a global scale, implementation is extremely variable. All mitigation is local, and the challenge is to provide the necessary resources to the communities that need them. Implementation requires two components: education and incentives. Education about the risk and available mitigation approaches is the first component. But, even when this information is provided, it can be difficult to motivate action for an event that may or may not occur within any individual's lifetime. Incentives are therefore also necessary and can be offered through legal mandate or economic benefit. As the population continues to grow in underdeveloped nations, where cities are increasingly concentrated in earthquake-prone locations and where current mitigation is least effective, the challenge to bridge the implementation gap could not be greater and of more importance.

Acknowledgments

Gilead Wurman provided several of the figures that were generated with GMT (Wessel and Smith, 1995). Support for this research was provided by USGS/NEHRP Grant # 06HQAG0147.

References

Aagaard B (2006) Finite-element simulations of ground motions in the San Francisco Bay Area from large earthquakes on the San Andreas Fault (abstract). *Seismological Research Letters* 77: 275.

Aki K (1980) Possibilities of seismology in the 1980s. *Bulletin of the Seismological Society of America* 70: 1969–1976.

Allen CR (1976) Responsibilities in earthquake prediction. *Bulletin of the Seismological Society of America* 66: 2069–2074.

Allen R (1978) Automatic earthquake recognition and timing from single traces. *Bulletin of the Seismological Society of America* 68: 1521–1532.

Allen R (1982) Automatic phase pickers: Their present use and future prospects. *Bulletin of the Seismological Society of America* 72: S225–S242.

Allen RM (2004) Rapid magnitude determination for earthquake early warning. In: Pecce M, Manfredi G, and Zollo A (eds.) *The Many Facets of Seismic Risk*, pp. 15–24. Napoli: Universita degli Studi di Napoli "Federico II".

Allen RM (2006) Probabilistic warning times for earthquake ground shaking in the San Francisco Bay Area. *Seismological Research Letters* 77: 371–376.

Allen RM (2007) The ElarmS earthquake warning methodology and application across California. In: Gasparini P and Zschau J (eds.) *Seismic Early Warning*, pp. 21–44. New York: Springer.

Allen RM and Kanamori H (2003) The potential for earthquake early warning in southern California. *Science* 300: 786–789.

Anderson J, Quaas R, Singh SK, et al. (1995) The Copala, Guerrero, Mexico earthquake of September 14, 1995 ($M_W = 7.4$): A preliminary report. *Seismological Research Letters* 66: 11–39.

Bakun WH (1999) Seismic activity of the San Francisco Bay Region. *Bulletin of the Seismological Society of America* 89: 764–784.

Bakun WH and Lindh AG (1985) The Parkfield, California, earthquake prediction experiment. *Science* 229: 619–624.

Bakun WH and McEvilly TV (1979) Earthquakes near Parkfield, California: Comparing the 1934 and 1966 sequences. *Science* 205: 1375–1377.

Berberian M (1990) Natural hazards and the first earthquake catalogue of Iran. *Historical Hazards in Iran Prior to 1900*, vol. 1, p. 649. Tehran: IIEES.

Bilham R (1988) Earthquakes and urban-growth. *Nature* 336: 625–626.

Bilham R (1996) Global fatalities from earthquakes in the past 2000 years: Prognosis for the next 30. In: Rundel JB, Turcotte DL, and Klein W (eds.) *Reduction and Predictability of Natural Disasters*, pp. 19–32. Reading, MA: Addison-Wesley.

Bilham R (1998) Death toll from earthquakes. *Geotimes* 43: 4.

Bilham R (2004) Urban earthquake fatalities: A safer world, or worse to come? *Seismological Research Letters* 75: 706–712.

Boatwright J and Bundock H (2005) Modified Mercalli Intensity Maps for the 1906 San Francisco earthquake plotted in ShakeMap Format. *Series: USGS Open File Report 2005-1135*.

Boese M, Erdik M, and Wenzel F (2004) Real-time prediction of ground motion from P-wave records. In: *EOS, Transactions of the American Geophysical Union Fall Meeting Supplement 85: Abstract S21A.0251*.

Boore DM and Zoback MD (1974) Two dimensional kinematic fault modeling of the Pacoima Dam strong motion recordings of February 9, 1971, San Fernando Earthquake. *Bulletin of the Seismological Society of America* 64: 555–570.

Brocher TM (2005) Compressional and shear wave velocity versus depth in the San Francisco Bay area, California: Rules for USGS Bay Area Velocity Model 05.0.0. *Series: USGS Open-File Report 2005-1317*.

Burridge R and Knopoff L (1967) Model and theoretical seismicity. *Bulletin of the Seismological Society of America* 57: 341–371.

Comerio MC, Tobriner S, and Fehrenkamp A (2006) Bracing Berkeley: A guide to seismic safety on the UC Berkeley campus. *Series, PEER 2006/01*Pacific Earthquake Engineering Research Center.

Cooper JD (1868) Earthquake indicator. *San Francisco Evening Bulletin* 10, 3 November 1868.

Cornell CA (1968) Engineering seismic risk analysis. *Bulletin of the Seismological Society of America* 58: 1583–1606.

Cua GB (2005) *Creating the Virtual Seismologist: Developments in Ground Motion Characterization and Seismic Early Warning*. PhD Thesis, Caltech.

d'Alessio MA, Johanson IA, Burgmann R, Schmidt DA, and Murray MH (2005) Slicing up the San Francisco Bay Area: Block kinematics and fault slip rates from GPS-derived surface velocities. *Journal of Geophysical Research* 110: B06403.

Dolenc D, Dreger D, and Larsen S (2006) 3D simulations of ground motions in northern California using the USGS SF06 velocity model (abstract). *Seismological Research Letters* 77: 300.

Dreger DS, Gee L, Lombard P, Murray MH, and Romanowicz B (2005) Rapid finite-source analysis and near-fault strong ground motions; application to the 2003 Mw 6.5 San Simeon and 2004 Mw 6.0 Parkfield earthquakes. *Seismological Research Letters* 76: 40–48.

Dreger D and Kaverina A (2000) Seismic remote sensing for the earthquake source process and near-source strong shaking: A case study of the October 16, 1999 hector mine earthquake. *Geophysical Research Letters* 27: 1941–1944.

Dunbar PK, Lockridge PA, and Whitewide LS (2006) Catalog of significant earthquakes 2150 B.C. to the present. *Series: Report SE-49*NOAA National Geophysical Data Center, http://www.ngdc.noaa.gov/nndc/struts/form?t=101650&s=1&d=1 (accessed January 2007).

Earle PS, Wald DJ, and Lastowka LA (2005) PAGER – Rapid assessment and notification of an earthquake's impact. *Series, USGS Fact Sheet 2005–2035*.

Erdik MO, Fahjan Y, Ozel O, Alcik H, Aydin M, and Gul M (2003) Istanbul earthquake early warning and rapid response system. In: *EOS, Transactions of the American Geophysical Union Fall Meeting Supplement 84: Abstract S42B.0153* .

Espinosa-Aranda JM, Jimenez A, Ibarrola G, et al. (1995) Mexico City seismic alert system. *Seismological Research Letters* 66: 42–53.

Fedotov SA (1965) Regularities of distribution of strong earthquakes in Kamchatka, the Kuril Islands and northern Japan (in Russian). *Akademiya Nauk SSSR, Institut Fiziki Zemli Trudi* 36: 66–93.

Gee LS, Neuhauser DS, Dreger DS, Pasyanos ME, Uhrhammer RA, and Romanowicz B (1996) Real-time seismology at UC Berkeley: The Rapid Earthquake Data Integration project. *Bulletin of the Seismological Society of America* 86: 936–945.

Gee L, Neuhauser D, Dreger D, Uhrhammer R, Romanowicz B, and Pasyanos M (2003) The rapid earthquake integration project. In: Lee WHK, Kanamori H, Jennings PC, and Kisslinger C (eds.) *International Handbook of Earthquake & Engineering Seismology*, pp. 1261–1273. San Diego: Academic Press.

Geller RJ (1996) Debate on evaluation of the VAN method: Editor's introduction. *Geophysical Research Letters* 23: 1291.

Geller RJ (1997) Earthquake prediction: A critical review. *Geophysical Journal International* 131: 425–450.

Geller RJ, Jackson DD, Kagan YY, and Mulargia F (1997) Geoscience – Earthquakes cannot be predicted. *Science* 275: 1616.

Giardini D (1999) The global seismic hazard assessment program (GSHAP) – 1992/1999. *Annali Di Geofisica* 42: 957–974.

Giardini D, Grunthal G, Shedlock KM, and Zhang PZ (1999) The GSHAP global seismic hazard map. *Annali Di Geofisica* 42: 1225–1230.

Gokhberg MB, Morgounov VA, and Pokhotelov OA (1995) *Earthquake Prediction: Seismo-Electromagnetic Phenomena*. Singapore: Gordon and Breach Publishers, 193 pp.

Grasso VF (2005) *Seismic Early Warning Systems: Procedure for Automated Decision Making*. PhD Thesis, Universita delgi Studi di Napoli Federico II.

Grasso VF and Allen RM (in review) Uncertainty in real-time earthquake hazard predictions. *Bulletin of the Seismological Society of America*.

Grunthal G, Bosse C, Sellami S, Mayer-Rosa D, and Giardini D (1999) Compilation of the GSHAP regional seismic hazard for Europe, Africa and the Middle East. *Annali Di Geofisica* 42: 1215–1224.

Gupta SK and Patwardham AM (1988) *Earthquake Prediction: Present Status*. Pune: University of Poona, 280 pp.

Hanks TC (1975) Strong ground motion of the San Fernando, California earthquake: Ground displacements. *Bulletin of the Seismological Society of America* 65: 193–225.

Hardebeck JL, Boatwright J, Dreger D, et al. (2004) Preliminary report on the 22 December 2003, M 6.5 San Simeon, California earthquake. *Seismological Research Letters* 75: 155–172.

Harris RA and Simpson RW (1998) Suppression of large earthquakes by stress shadows: A comparison of Coulomb and rate-and-state failure. *Journal of Geophysical Research* 103: 24439–24451.

Hauksson E, Jones LM, and Shakal AF (2003) TriNet: A modern ground-motion seismic network. In: Lee WHK, Kanamori H, Jennings PC, and Kisslinger C (eds.) *International Handbook of Earthquake & Engineering Seismology*, pp. 1276–1284. San Diego: Academic Press.

Hauksson E, Small P, Hafner K, et al. (2001) Southern California Seismic Network: Caltech/USGS element of TriNet 1997–2001. *Seismological Research Letters* 72: 690–704.

Heaton TH (1985) A model for a seismic computerized alert network. *Science* 228: 987–990.

Horiuchi S, Negishi H, Abe K, Kamimura A, and Fujinawa Y (2005) An automatic processing system for broadcasting earthquake alarms. *Bulletin of the Seismological Society of America* 95: 708–718.

Housner GW (1984) Historical view of earthquake engineering. In: *Proceedings of the Eighth World Conference on Earthquake Engineering, San Francisco,* pp. 25–39.

Housner GW, Bergman LA, Caughey TK, et al. (1997) Structural control: Past, present, and future. *Journal of Engineering Mechanics, ASCE* 123: 897–971.

Housner GW, Martel R, and Alford JL (1953) Spectrum analysis of strong motion earthquakes. *Bulletin of the Seismological Society of America* 42: 97–120.

Hudson DE (1972) Local distribution of strong earthquake ground motions. *Bulletin of the Seismological Society of America* 62: 1765–1786.

Ikeya M (2004) *Earthquakes and Animals*. Singapore: World Scientific Publishing Company, 295 pp.

Isikara AM and Vogel A (1982) *Multidisciplinary Approach to Earthquake Prediction*. Braunschweig: Friodr. Vieweg & Sohn, 578 pp.

Jackson DD and Kagan YY (1993) Comment on "Seismic gap hypothesis: Ten years after", Reply to S.P. Nishenko and L.R. Sykes. *Journal of Geophysical Research* 98: 9917–9920.

Ji C, Helmberger DV, and Wald DJ (2004) A teleseismic study of the 2002 Denali fault, Alaska, earthquake and implications for rapid strong-motion estimation. *Earthquake Spectra* 20: 617–637.

Kagan YY and Jackson DD (1991) Seismic gap hypothesis. *Journal of Geophysical Research* 96: 21419–21431.

Kagan YY and Jackson DD (1995) New seismic gap hypothesis: Five years after. *Journal of Geophysical Research* 100: 3943–3959.

Kamigaichi O (2004). *Journal of Japan Association for Earthquake Engineering* 4.

Kanamori H (2003) Earthquake prediction: An overview. In: Lee WHK, Kanamori H, Jennings PC, and Kisslinger C (eds.) *International Handbook of Earthquake & Engineering Seismology*, pp. 1205–1216. San Diego: Academic Press.

Kanamori H (2005) Real-time seismology and earthquake damage mitigation. *Annual Review of Earth and Planetary Sciences* 33: 195–214.

Kanamori H, Hauksson E, and Heaton T (1991) TERRAscope and CUBE project at Caltech. *Eos, Transactions American Geophysical Union* 72: 564.

Kaverina A, Dreger D, and Price E (2002) The combined inversion of seismic and geodetic data for the source process of the 16 October 1999 M-w 7.1 Hector Mine, California, earthquake. *Bulletin of the Seismological Society of America* 92: 1266–1280.

Kelleher JA, Sykes LR, and Oliver J (1973) Criteria for prediction of earthquake locations, Pacific and Caribbean. *Journal of Geophysical Research* 78: 2547–2585.

Kinoshita S (2003) Kyoshin Net (K-Net), Japan. In: Lee WHK, Kanamori H, Jennings PC, and Kisslinger C (eds.) *International Handbook of Earthquake & Engineering Seismology*, pp. 1049–1056. San Diego: Academic Press.

Kircher CA, Seligson HA, Bouabid J, and Morrow GC (2006) When the big one strikes again – Estimated losses due to a repeat of the 1906 San Francisco earthquake. *Earthquake Spectra* 22: S297–S339.

Knopoff L (1996) Earthquake prediction: The scientific challenge. *Proceedings of the National Academy of Sciences of the United States of America* 93: 3719–3720.

Kohler M, Magistrale H, and Clayton R (2003) Mantle heterogeneities and the SCEC three-dimensional seismic velocity model version 3. *Bulletin of the Seismological Society of America* 93: 757–774.

Krishnan S, Ji C, Komatitsch D, and Tromp J (2006) Case studies of damage to tall steel moment-frame buildings in southern California during large San Andreas earthquakes. *Bulletin of the Seismological Society of America* 96: 1523–1537.

Langbein J, Borcherdt R, and Dreger D (2005) Preliminary report on the 28 September 2004, M 6.0 Parkfield, California Earthquake. *Seismological Research Letters* 76: 10–26.

Lawson AC (1908) *The California Earthquake of April 18, 1906*. Washington, DC: Carnegie Institution of Washington.

Lee WHK, Shin TC, Kuo KW, Chen KC, and Wu CF (2001) CWB free-field strong-motion data from the 21 September Chi-Chi, Taiwan, Earthquake. *Bulletin of the Seismological Society of America* 91: 1370–1376.

Iervolino I, Convertito V, Giorgio M, Manfredi G, and Zollo A (2006) Real-time risk analysis for hybrid earthquake early warning systems. *Journal of Earthquake Engineering* 10: 867–885.

Lockman A and Allen RM (2005) Single station earthquake characterization for early warning. *Bulletin of the Seismological Society of America* 95: 2029–2039.

Lockman A and Allen RM (2007) Magnitude-period scaling relations for Japan and the Pacific Northwest: Implications for earthquake early warning. *Bulletin of the Seismological Society of America* 97: 140–150.

Lomnitz C (1994) *Fundamentals of Earthquake Prediction*. New York: John Wiley & Sons, 326 pp.

McCann WR, Nishenko SP, Sykes LR, and Krause J (1979) Seismic gaps and plate tectonics: Seismic potential for major boundaries. *Journal of Pure and Applied Geophysics* 117: 1082–1147.

McCue K (1999) Seismic hazard mapping in Australia, the southwest Pacific and southeast Asia. *Annali Di Geofisica* 42: 1191–1198.

McGuire RK (1993) Computation of seismic hazard. *Annali Di Geofisica* 34: 181–200.

Milne J and Burton WK (1891) *The Great Earthquake in Japan, 1891, 2nd edn., Lane,* 69 pp. Japan: Crawford & Co, Yokohama, and 30 plates.

Mogi K (1985) *Earthquake Prediction*. London: Academic Press, 335 pp.

Mori J, Kanamori H, Davis J, and Hauksson E (1998) Major improvements in progress for southern California earthquake monitoring. *Eos* 79: 217–221.

Nakamura Y (1988) On the urgent earthquake detection and alarm system (UrEDAS). In: *Proceedings of the 9th World Conference on Earthquake Engineering VII,* pp. 673–678.

Nakamura Y (2004) UrEDAS, urgent earthquake detection and alarm system, now and future. In: *Proceedings of the 13th World Conference on Earthquake Engineering, August 2004,* Paper No. 908.

Nakamura Y and Tucker BE (1988) Earthquake warning system for Japan Railways' Bullet Trains: Implications for disaster prevention in California. *Earthquakes and Volcanoes* 20: 140–155.

National Research Council (1971) *The San Fernando Earthquake of February 9, 1971: Lessons from a Moderate Earthquake on the Fringe of a Densely Populated Region*. Washington, DC: National Academy Press, 24 pp.

National Research Council (2002) *Living on an Active Earth: Perspectives on Earthquake Science*. Washington, DC: National Academies Press, 418 pp.

National Research Council (2003) *Preventing Earthquake Disasters: The Grand Challenge in Earthquake Engineering*. Washington, DC: National Academy Press, 172 pp.

Nishenko SP (1989) Earthquakes, hazards and predictions. In: James DE (ed.) *The Encyclopedia of Solid-Earth Geophysics*, pp. 260–268. New York: Van Nostrand Reinhold.

Nishenko SP (1991) Circum-Pacific seismic potential: 1989–1999. *Journal of Pure and Applied Geophysics* 135: 169–259.

Nishenko SP and Sykes LR (1993) Comment on "Seismic gap hypothesis: Ten years after" by Y.Y. Kagan and D.D. Jackson. *Journal of Geophysical Research* 98: 9909–9916.

Odaka T, Ashiya K, Tsukada S, Sato S, Ohtake K, and Nozaka D (2003) A new method of quickly estimating epicentral distance and magnitude from a single seismic record. *Bulletin of the Seismological Society of America* 93: 526–532.

Office of Statewide Health Planning and Development (2001) *Summary of Hospital Performance Ratings, Series*.

Office of Technology Assessment (1995) *Reducing Earthquake Losses*. Washington, DC: U.S. Government Printing Office.

Olsen KB, Day SM, and Minster JB (2006) Strong shaking in Los Angeles expected from southern San Andreas earthquake. *Geophysical Research Letters* 33: L07305.

Olson E and Allen RM (2005) The deterministic nature of earthquake rupture. *Nature* 438: 212–215.

Olson EL and Allen RM (2006) Is earthquake rupture deterministic? (Reply). *Nature* 442: E6.

Olson RS, Podesta B, and Nigg JM (1989) *The Politics of Earthquake Prediction*. Princeton: Princeton University Press, 187 pp.

Otsuka M (1972) A chain-reaction-type source model as a tool to interpret the magnitude-frequency relation of earthquakes. *Journal of Physics of the Earth* 20: 35–45.

Perkins J (2003) *San Francisco and the Bay Area Earthquake Nightmare, Series*Association of Bay Area Governments.

Reid HF (1910) *The Mechanics of the Earthquake*. Washington, DC: Carnegie Institution of Washington, 192 pp.

Rikitake T (1976) *Earthquake Prediction*. Amsterdam: Elsevier, 357 pp.

Rikitake T (1982) *Earthquake Forecasting and Warning*. Dordrecht: D. Reidel Publishing, 402 pp.

Rikitake T (1986) *Earthquake Premonitory Phenomena: Database for Earthquake Prediction*. Tokyo: Tokyo University Press, 232 pp.

Rikitake T and Hamada K (2001) Earthquake prediction. In: Meyers RA (ed.) *Encyclopedia of Physical Science and Technology*, 3rd edn., pp. 743–760. San Diego: Academic Press.

RMS (1995) What if the 1906 earthquake strikes again? A San Francisco Bay Area scenario. *Series: RMS Report:* Risk Management Solutions, Inc.

Romanowicz B (1993) Spatiotemporal patterns in the energy release of great earthquakes. *Science* 260: 1923–1926.

Rowshandel B (2006) Estimation of future earthquake losses in California. In: Loyd R, Mattison E, and Wilson R (eds.) *Earthquakes of the San Francisco Bay Area and Northern California*. Sacramento: California Geological Survey.

Rydelek P and Horiuchi S (2006) Is earthquake rupture deterministic? *Nature* 442: E5–E6.

Shedlock KM and Tanner JG (1999) Seismic hazard map of the western hemisphere. *Annali Di Geofisica* 42: 1199–1214.

Simons FJ, Dando B, and Allen RM (2007) Automatic detection and rapid determination of earthquake magnitude by wavelet multiscale analysis of the primary arrival. *Earth and Planetary Science Letters* 250: 214–223.

Sobolev GA (1995) *Fundamentals of Earthquake Prediction*. Moscow: Electromagnetic Research Center, 161 pp.

Somerville P, Irikura K, Graves RP, Sawada S, and Wald D (1999) Characterizing crustal earthquake slip models for the prediction of strong motion. *Seismological Research Letters* 70: 59–80.

Song S, Beroza G, and Segall P (2006) A unified source model for the 1906 San Francisco Earthquake (abstract). *Seismological Research Letters* 77: 271.

Sorensen SP and Meyer KJ (2003) Effect of the Denali Fault rupture on the Trans-Alaska Pipeline. In: *Proceedings of the 6th U.S. Conference on Lifeline Earthquake Engineering, Long Beach, CA, August*, 1–9.

Sykes LR, Shaw BE, and Scholz CH (1999) Rethinking earthquake prediction. *Journal of Pure and Applied Geophysics* 155: 207–232.

Tucker BE (2004) Trends in global urban earthquake risk: A call to the international earth science and earthquake engineering communities. *Seismological Research Letters* 75: 695–700.

Turcotte DL (1992) *Fractals and Chaos in Geology and Geophysics*. Cambridge: Cambridge University Press, 221 pp.

Unesco (1984) *Earthquake Prediction*. Tokyo: Terra Scientific Publishing Company, 995 pp.

United Nations (2004) *World Urbanization Prospects. The 2003 Revision*. New York: United Nations, 185 pp.

Veneziano D, Cornell CA, and O-Hara T (1984) Historical method of seismic hazard analysis. *Series, NP-3438*.

Vogel A (1979) *Terrestrial and Space Techniques in Earthquake Prediction Research*. Braunschweig: Friodt. Vieweg & Sohn, 712 pp.

Wald DJ, Earle PS, Lin K-W, Quitoriano V, and Worden B (2006a) Challenges in rapid ground motion estimation for the prompt assessment of global urban earthquakes. In: *Proceedings of the 2nd International Workshop on Strong Ground Motion Prediction and Earthquake Tectonics in Urban Areas, ERI*. Tokyo: BERI, 8 pp.

Wald DJ, Quitoriano V, and Dewey JW (2006b) USGS "Did you feel it?" Community internet intensity maps: Macroseismic data collection via the internet. In: *ISEE Proceedings, Geneva*.

Wald DJ, Quitoriano V, Heaton TH, and Kanamori H (1999a) Relationships between peak ground acceleration, peak ground velocity, and Modified Mercalli Intensity in California. *Earthquake Spectra* 15: 557–564.

Wald DJ, Quitoriano V, Heaton TH, Kanamori H, Scrivner CW, and Worden CB (1999b) TriNet "ShakeMaps": Rapid generation of peak ground motion and intensity maps for earthquakes in southern California. *Earthquake Spectra* 15: 537–555.

Wald DJ, Worden CB, Quitoriano V, and Pankow KL (2005) ShakeMap manual: Technical manual, users guide, and software guide. *Series, Techniques and Methods 12–A1:* U.S. Geological Survey.

Wang K, Chen Q-F, Sun S, and Wang A (2006) Predicting the 1975 Haicheng earthquake. *Bulletin of the Seismological Society of America* 96: 757–795.

Wessel P and Smith WHF (1995) New version of the Generic Mapping Tools released. *Supplement EOS Transactions, American Geophysical Union* 76: 329.

WG02 (2003) Earthquake probabilities in the San Francisco Bay Region: 2002–2031. *Series: Open File Report 03-214*. US Geological Survey.

Whittaker A, Moehle J, and Higashino M (1998) Evolution of seismic design practice in Japan. *Structural Design of Tall Buildings* 7: 93–111.

Wu Y-M and Kanamori H (2005a) Experiment on an onsite early warning method for the Taiwan early warning system. *Bulletin of the Seismological Society of America* 95: 347–353.

Wu YM and Kanamori H (2005b) Rapid assessment of damage potential of earthquakes in Taiwan from the beginning of P waves. *Bulletin of the Seismological Society of America* 95: 1181–1185.

Wu Y-M, Kanamori H, Allen RM, and Hauksson E (2007) Experiment using the tau-c and Pd method for earthquake early warning in southern California. *Geophysical Journal International* 170: 711–717.

Wu YM, Shin TC, and Tsai YB (1998) Quick and reliable determination of magnitude for seismic early warning. *Bulletin of the Seismological Society of America* 88: 1254–1259.

Wu Y-M and Teng T-I (2002) A virtual subnetwork approach to earthquake early warning. *Bulletin of the Seismological Society of America* 92: 2008–2018.

Wu Y-M, Yen H-Y, Zhao L, Huang B-S, and Liang W-T (2006) Magnitude determination using initial P waves: A single-station approach. *Geophysical Research Letters* 33: L05306.

Wu YM and Zhao L (2006) Magnitude estimation using the first three seconds P-wave amplitude in earthquake early warning. *Geophysical Research Letters* 33: L16312.

Wurman G, Allen RM, and Lombard P (2007) Toward earthquake early warning in northern California. *Journal of Geophysical Research* 112(B08311): 1–19.

Wyss M (1979) *Earthquake Prediction and Seismicity Patterns*. (Reprinted from Pure and Applied Geophysics, Vol 117, No 6, 1979). Basel: Birkhauser Verlag, 237 pp.

Wyss M (1991) *Evaluation of Proposed Earthquake Precursors*. Washington, DC: American Geophysical Union, 94 pp.

Wyss M (2004) Human losses expected in Himalayan earthquakes. *Natural Hazards* 34: 305–314.

Yamada M and Heaton T (2006) Early warning systems for large earthquakes: Extending the virtual seismologist to finite ruptures (abstract). *Seismological Research Letters* 77: 313.

Yamakawa K (1998) The Prime Minister and the earthquake: Emergency management leadership of Prime Minister Marayama on the occasion of the Great Hanshin-Awaji earthquake disaster. *Kansai University Review of Law and Politics* 19: 13–55.

Zhang P, Yang Z-X, Gupta HK, Bhatia SC, and Shedlock KM (1999) Global seismic hazard assessment program (GSHAP) in continental Asia. *Annali Di Geofisica* 42: 1167–1190.

This page intentionally left blank

4.24 Complexity and Earthquakes

R Shcherbakov, Western University, London, ON, Canada
DL Turcotte and JB Rundle, University of California, Davis, CA, USA

© 2015 Elsevier B.V. All rights reserved.

4.24.1	**Introduction**	628
4.24.2	**Relevant Observations**	629
4.24.2.1	GR Frequency–Magnitude Scaling	629
4.24.2.2	Båth's Law	631
4.24.2.3	The Omori Law	632
4.24.2.4	Accelerated Moment Release	635
4.24.2.5	Recurrence Times	637
4.24.2.6	Interoccurrence Times	638
4.24.3	**Self-Organizing Complex Systems**	640
4.24.3.1	Fractality	640
4.24.3.2	Deterministic Chaos	640
4.24.3.3	Criticality and SOC	640
4.24.4	**Relevant Models**	642
4.24.4.1	Cellular Automaton Models	642
4.24.4.2	Fiber-Bundle Models	643
4.24.4.3	Numerical Simulations	644
4.24.4.4	Damage Mechanics Models	645
4.24.4.5	Epidemic-Type Aftershock Sequence Models	647
4.24.5	**Discussion**	647
4.24.5.1	Overview	647
4.24.5.2	Earthquake Forecasting (Prediction)	647
4.24.5.3	Future Directions	648
Acknowledgments		649
References		649

Glossary

Accelerated moment release (AMR) A systematic regional increase in seismic activity prior to a large earthquake.

Aftershocks Earthquakes that follow the largest earthquake in a sequence.

Båth's law An empirical observation that the difference in magnitude between the mainshock and the largest aftershock is approximately constant.

Cellular automaton A model consisting of a grid of elements that interact and evolve in time through specified rules.

Characteristic earthquakes A sequence of earthquakes that occur quasiperiodically on a fault or fault segment with approximately equal magnitudes.

Correlation length A characteristic length scale (spatial or temporal) over which a point disturbance (event) can influence the system.

Critical point A point in the phase space of a statistical mechanics system where the first derivative of thermodynamic functions is continuous and the correlation length tends to infinity. In the vicinity of the critical point, thermodynamic functions behave as power laws with corresponding critical exponents.

Damage mechanics A branch of continuum mechanics that studies nonelastic irreversible deformation of solids.

Foreshocks Earthquakes that precede the largest earthquake in a sequence.

Fractal A geometrical structure with a fractional dimension. It is also possible to describe a power-law dependence of the number of events on the characteristic length of the events as defined in eqn [40].

Gutenberg–Richter scaling A relation between the cumulative number of earthquakes and their magnitudes as defined in eqn [1].

Interoccurrence times The time intervals between successive earthquakes in a region.

Mainshock The largest earthquake in an associated sequence of earthquakes.

Recurrence times The time intervals between characteristic earthquakes.

Self-organized criticality (SOC) A theory describing the behavior of nonlinear, driven, threshold systems consisting of many interacting elements. The self-organized critical system oscillates about a quasisteady critical state, and there are a steady input of energy into the system and a stochastic output through avalanches whose magnitudes are distributed according to a power law. The critical state is

reached through internal dynamics without any tuning of model parameters.

The Omori law An empirical law describing the dependence of the rate of aftershock occurrence on the time since the mainshock as defined in eqn [7].

Nomenclature

A	Area (m^2)
b	b-Value
c	Characteristic time (s)
C_v	Coefficient of variation
D	Fractal dimension
E	Seismic energy release (J)
E	Young's modulus (Pa)
e	Energy density ($J\ kg^{-1}$)
k_c	Connecting spring constant ($N\ m^{-1}$)
k_L	Loader plate spring constant ($N\ m^{-1}$)
M	Seismic moment (J)
m	Earthquake magnitude
N	Number of events
N_{as}	Number of aftershocks
P	Cumulative distribution function
p	Exponent in the modified Omori law
r	Range of interaction (m)
r	Decay rate of aftershocks
s	Power-law exponent
t	Time (s)
t_d	Characteristic damage time (s)
t_f	Time to failure (s)
v	Hazard rate (s^{-1})
V	Driver plate velocity ($m\ s^{-1}$)

Symbols

α	Damage variable
β	Modified b-value
μ	Mean
ξ	Correlation length (m)
ρ	Power-law exponent
σ	Standard deviation
σ	Stress (Pa)
σ_y	Yield stress (Pa)
τ	Characteristic time (s)
ε	Strain
ε_B	Cumulative Benioff strain
ε_y	Yield strain

4.24.1 Introduction

Complex behavior manifests itself in various natural phenomena. Examples can be found at different levels of organization of natural systems ranging from molecules and microorganisms to the large-scale structure of the universe and the existence of life. All these systems are characterized by highly nonlinear interactions among their constituent parts and behavior far from equilibrium. As a result, complexity is achieved through mechanisms of emergence, information transfer, and nonlinearity. It appears that nonlinearity plays an essential role in the emergence of complex behavior and is ubiquitous in natural systems (Nicolis and Nicolis, 2012).

Earthquakes represent one example of such complex behavior. They usually occur in the upper brittle layer of the Earth's crust as a result of instabilities generated by moving plates. Constant driving assisted by the convective movement of the mantle supplies the energy needed to sustain a seismogenic region in a state far from equilibrium. This state is reached through the processes of self-organization and is characterized by long-range temporal and spatial correlations and scale-invariant behaviors. It is also believed that correlations are inherently present in the Earth's crust and are signatures of the nonlinear dynamics of rheological, frictional, chemical, and other processes. From this point of view, the Earth's crust can be considered as a strongly correlated nonlinear dissipative system. A review of the collective behavior of earthquakes and faults has been given by Ben-Zion (2008).

Earthquake complexity is evident through various phenomenological observations, which are quantified using statistical and geophysical methods. The occurrence of earthquakes can be envisioned as a cascade of triggered events with scale-invariant features. Seismogenic zones where earthquakes strike encompass the network of faults embedded into a heterogeneous medium. Fault sizes themselves form (multi)fractal structures. The seismogenic zones evolve on long geological timescales reshaping the fault networks and creating new structures. They operate at different timescales and involve interaction among various constituent parts through physical, chemical, and rheological processes. The emergence of such complex behavior is a direct result of nonlinear, threshold dynamics and the dissipative nature of earthquake systems.

Another important aspect of the complexity of earthquakes is related to their forecasting (prediction). This becomes important as seismic activity in many cases is punctuated by the occurrence of large earthquakes with devastating consequences. The problem of forecasting is one of the signatures of complex systems. Nonlinearity, long-range spatial and temporal correlations, effects of memory, and feedback mechanisms lead to a multiplicity of states and a strong sensitivity to initial conditions. This results in the coexistence of order and disorder in the behavior of complex systems. This in turn

makes the predictability in the behavior of complex systems more difficult and in some cases almost impossible. This is directly applicable to earthquakes. However, within this complexity, one hopes to find robust probabilistic approaches in order to quantify and mitigate the earthquake hazard.

Theoretical studies of seismology are usually dominated by applications of classical physics. However, many aspects of seismicity exhibit complex nonlinear behaviors. These aspects are best addressed using concepts developed in modern statistical physics. Examples include scaling laws associated with phase changes and the chaotic behavior of dynamical systems. Related concepts include scale invariance and fractality, deterministic chaos in low-dimensional systems, and self-organized criticality (SOC). Relevant models are generally nonlinear and stochastic in nature, and the standard approaches are not applicable. Instead, numerical simulations are used.

To understand the complex behavior of seismicity, one has to apply the methods and approaches developed in nonequilibrium statistical physics to treat earthquakes as nonlinear, driven, dissipative threshold systems. Recent progress in this direction has been considered by a number of authors (Ben-Zion, 2008; Dahmen and Ben-Zion, 2009; Kanamori and Brodsky, 2004; Kawamura et al., 2012; Main, 1996; Rundle et al., 2003; Sornette, 2006; Sornette and Werner, 2009; Zöller et al., 2009). This approach revolves around the general idea that one has to understand the microscopic behavior of complex systems based on the fundamental physical principles that govern such systems. In addition, one has to take into account the processes of emergence, which are observed in complex systems where collective behavior of constituent parts gives rise to new phenomena, which cannot be explained solely by analyzing the microdynamics and structure of individual parts.

The main aim of this chapter is to introduce the reader into different aspects of scaling properties of earthquakes that are the manifestation of the complex nature of earthquake processes and illustrate the methods and techniques developed in statistical physics to study these properties. This chapter is organized as follows. In **Section 4.24.2**, we will discuss relevant observations concerning the occurrence and statistics of earthquakes. We will emphasize applicable empirical laws. Examples include the following:

1. Gutenberg–Richter (GR) frequency–magnitude relation (Gutenberg and Richter, 1954). This is a robustly applicable scaling law that relates the cumulative number of earthquakes to their magnitudes and rupture areas.
2. The Omori law for the decay of aftershock activity following a mainshock (Scholz, 2002; Utsu et al., 1995). This is another scaling law that gives the dependence of the rate of aftershock production on the time since the mainshock occurred.
3. Båth's law (Båth, 1965). This empirical law states that the magnitude of the largest detected aftershock has a constant difference relative to the magnitude of the mainshock.

Other less accepted, and in some cases more controversial, scaling laws include the following:

1. Accelerated moment release (AMR) prior to large earthquakes (Bufe and Varnes, 1993; Jaumé and Sykes, 1999; Knopoff, 1996; Sykes and Jaumé, 1990). This is a systematic increase in the cumulative moment (energy) of smaller and moderate earthquakes prior to a large earthquake.
2. Recurrence times. This is a statistical distribution of the recurrence times between characteristic earthquakes on a fault, possibly satisfying the Weibull distribution.
3. Interoccurrence statistics in both time and space of all earthquakes in a region. For aftershock sequences following large mainshocks, the distribution of interoccurrence times between aftershocks appears to satisfy the nonhomogeneous Poisson statistics to a good approximation.

Our presentations will emphasize aftershocks since the available data can be of very good quality (Kisslinger, 1996). The major problem with earthquakes in general is that the recurrence times are large compared with the period for which data are available. One approach used to overcome this problem is to utilize numerical simulations to produce synthetic earthquake catalogs.

In **Section 4.24.3**, we will discuss the general characteristics of self-organizing complex systems. These include the following:

1. Fractality, the applicability of power laws and self-similar scaling laws.
2. Deterministic chaos. Low-order systems related to seismicity have been shown to exhibit deterministic chaos. The implications of this association will be discussed.
3. Criticality and SOC. The applicability of these concepts is certainly controversial but must be addressed.

In **Section 4.24.4**, relevant models will be discussed. We classify these models as follows:

1. Cellular automaton models. The classic example relevant to seismicity is the multiple slider-block model.
2. Numerical simulations. Numerical simulations are widely used to forecast weather. Initial attempts to apply these approaches to seismicity are being carried out.
3. Damage mechanics models. Damage mechanics is used to model the complex brittle behavior in engineering materials. Applications to seismicity will be discussed.

Finally, in **Section 4.24.5**, the material presented earlier will be summarized. Future directions will also be considered. A specific example is earthquake forecasting (prediction).

4.24.2 Relevant Observations

4.24.2.1 GR Frequency–Magnitude Scaling

Earthquakes are known to universally (with a few exceptions) satisfy a relatively simple frequency–magnitude scaling. In terms of the earthquake magnitude m, this scaling satisfies the GR frequency–magnitude relation (Gutenberg and Richter, 1954):

$$\log N(\geq m) = a - bm \qquad [1]$$

where N is the cumulative number of earthquakes with a magnitude greater than m occurring in a specified area and time and b and a are constants. This relation is valid for earthquakes both regionally and globally. The constant b or 'b-value' varies from region to region but is generally in the

range of $0.8 < b < 1.2$ (Frohlich and Davis, 1993). Although the earthquake magnitude m has an empirical basis, we will show that eqn [1] is a power-law (fractal) relation. The constant a in eqn [1] is a measure of the regional intensity of the seismicity.

The frequency–magnitude distributions of the regional seismicity in southern California on a yearly basis are plotted in Figure 1 using data obtained from the *SCSN* catalog (http://www.data.scec.org). For each individual year between 1989 and 2012, the cumulative number of earthquakes $N(\geq m)$ with magnitudes greater than m is plotted as a function of local magnitude m. For the whole period 1989–2012 taken together the parameters of the GR power-law relation, eqn [1], per year were computed with $b = 1.00\backslash\text{pm } 0.02$ and $a = 5.60\backslash\text{pm } 0.12$ and the corresponding scaling is shown as the solid straight lines in Figure 1(a)–1(d). The estimation of the parameters was performed for earthquakes above $m = 3.0$. In Figure 1, there is generally good agreement between each individual year's data and the GR relation (solid straight line) for the period 1989–2012. The exceptions can be attributed to the aftershock sequences of the 1992 Landers, 1994 Northridge, 1999 Hector Mine, and 2010 El Mayor earthquakes.

For years without large numbers of aftershocks, the background seismicity in southern California illustrated in Figure 1 is nearly uniform from year to year and is not a function of time. Small earthquakes behave like thermal background noise. There is observational evidence that the Earth's crust is continuously on the brink of failure (Scholz, 1991). One example is induced seismicity. Whenever the crust is loaded, earthquakes are induced whether in a tectonically active area or not. Examples of nontectonic loading include the filling of a reservoir behind a newly completed dam or the high-pressure injection of fluids in a deep well. Induced seismicity has also been associated with hydraulic fracturing (fracking) during the extraction of shale gas and oil.

Studies have shown that aftershocks also satisfy the GR relation [1] with 'b-values' that are not statistically different

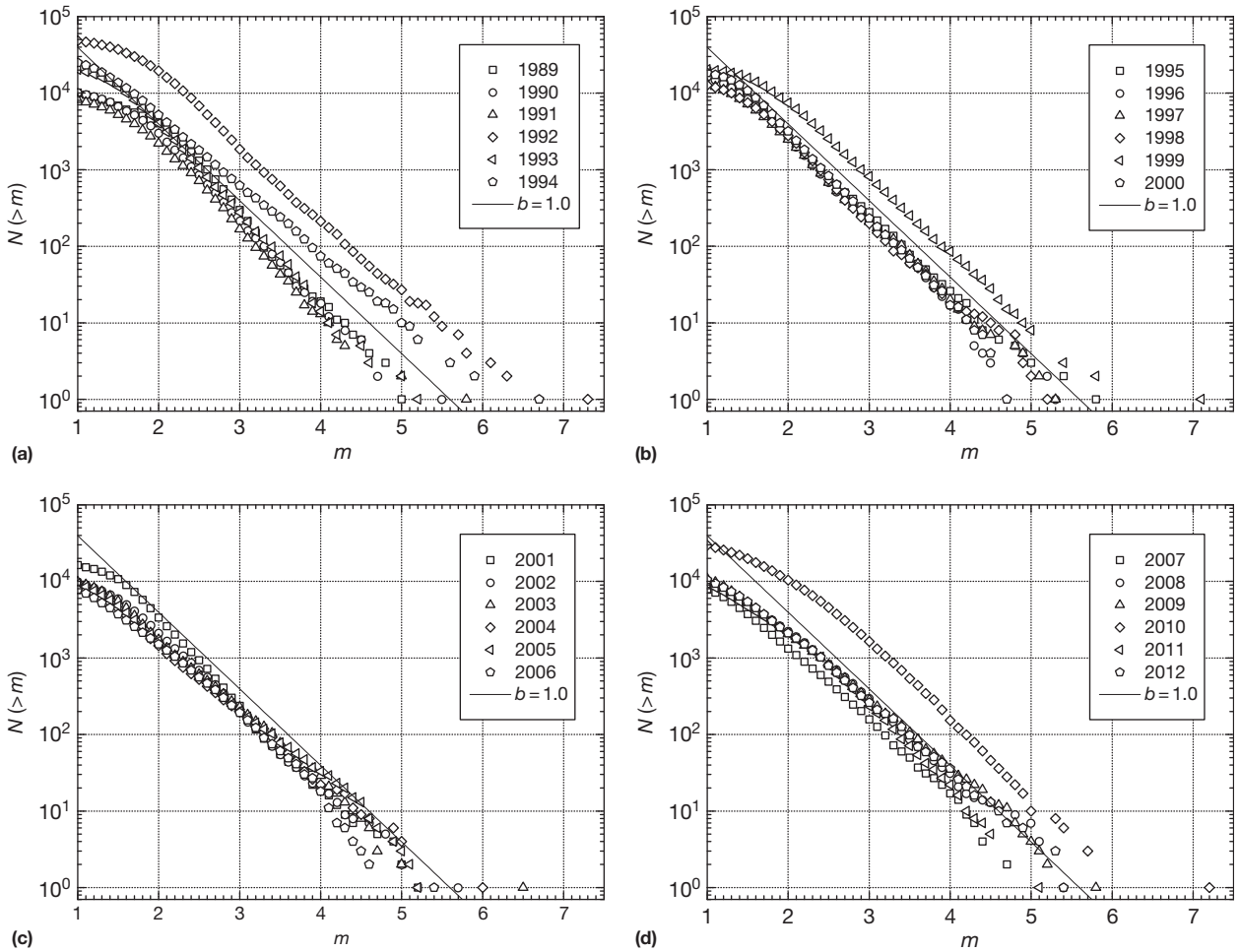


Figure 1 Cumulative number of earthquakes per year, $N(\geq m)$, occurring in southern California with magnitudes greater that m as a function of the local magnitude m. Twenty-four individual years are considered (SCSN catalog, http://www.data.scec.org): (a) 1989–1994, (b) 1995–2000, (c) 2001–2006, and (d) 2007–2012. The solid straight line in (a)–(d) is the Gutenberg–Richter relation, eqn [1], with $b = 1.00 \pm 0.02$ and $a = 5.60 \pm 0.12$ estimated for magnitudes above $m = 3.0$. The larger number of earthquakes in 1992, 1994, 1999, and 2010 can be attributed to the aftershocks of the Landers, Northridge, Hector Mine, and El Mayor earthquakes, respectively. When aftershocks are excluded, the background seismicity in southern California is nearly uniform in time.

from the frequency–magnitude distribution of all earthquakes (Kisslinger, 1996). The Parkfield earthquake (28 September 2004; $m_{ms} = 6.0$; epicenter 35.818° N, −120.366° W) generated a well-documented aftershock sequence (Bakun et al., 2005; Shcherbakov et al., 2006a). We will consider as aftershocks all earthquakes that occurred in an elliptical region centered at 35.9° N and −120.5° W with the radii of sizes 0.4° and 0.15° and oriented 137° NW during the 365-day period following the earthquake. The cumulative numbers of aftershocks with magnitudes greater than m, $N(>m)$, are given as a function of m in Figure 2. The maximum likelihood fit of eqn [1] to the Parkfield data is obtained by taking $b = 0.92 \pm 0.16$ and $a = 4.43 \pm 0.39$ for all aftershocks above $m = 2.5$.

Presumably, aftershocks satisfy the GR scaling for the same reason that all earthquakes do. However, no generally accepted theory is available for the explanation of scale-invariant nature of this distribution. Generally, two end-member models have been proposed: (1) Each fault has a GR distribution of earthquakes. The largest earthquake on the fault would be limited by the size of the fault, but earthquakes occurring on the fault would satisfy GR scaling. (2) There is a power-law frequency–area distribution of faults, and each fault has a reoccurrence of characteristic earthquakes. Large faults have large characteristic earthquakes of approximately equal size that rupture the entire fault; similarly, small faults have small characteristic earthquakes. It is recognized that faults have a fractal number–area distribution (Turcotte, 1997), and this distribution is responsible for the GR scaling of regional seismicity. The validity of the end-member models has been discussed by Wesnousky (1994) and more recently by Yoder et al. (2012b). It is agreed that actual seismicity will not satisfy either extreme model but will have characteristics of both. However, the importance of characteristic earthquakes is generally accepted as will be discussed in Section 4.24.2.5.

An alternative approach has been suggested that the observed scale-invariant distribution of earthquake magnitudes is due to the self-organized dynamics of seismogenic zones (Bak and Tang, 1989; Turcotte, 1999b). In this approach, the internal nonlinear dynamics of the upper part of the crust is responsible for the emergence of the scale-invariant behavior. Several models displaying SOC have been introduced to illustrate and confirm this mechanism and will be discussed in Section 4.24.4. It has been also suggested that the seismogenic crust can be considered as a statistical mechanics system operating near the critical point (Sammis and Sornette, 2002) or spinodal (Rundle et al., 2003). In this picture, background seismicity drives the fault system toward criticality while large earthquakes destroy long-range correlations and push the system away from a critical point.

Another important question is the range of validity of GR scaling. On a global basis, the validity is established for magnitudes less than $m \leq 7.6$. Above this value, there may be transition to higher b-value, $b \approx 1.5$ (Rundle, 1989a; Yoder et al., 2012a), or a gradual rollover (Burroughs and Tebbens, 2002; Sornette et al., 1996). It is established that GR scaling is generally valid for $m > 0$ and, based on studies in deep mines, may be valid for $m > -4.3$ (Kwiatek et al., 2010).

4.24.2.2 Båth's Law

Another important empirical law concerning aftershocks is Båth's law. This law states that the difference in magnitude between the mainshock and its largest aftershock is approximately constant independent of the magnitude of the mainshock (Båth, 1965). That is,

$$\Delta m = m_{ms} - m_{as}^{max} \qquad [2]$$

where m_{ms} is the magnitude of the mainshock, m_{as}^{max} is the magnitude of the largest aftershock, and typically $\Delta m \approx 1.2$.

A number of extensive studies of the statistical variability of Δm have been carried out (Console et al., 2003a; Felzer et al.,

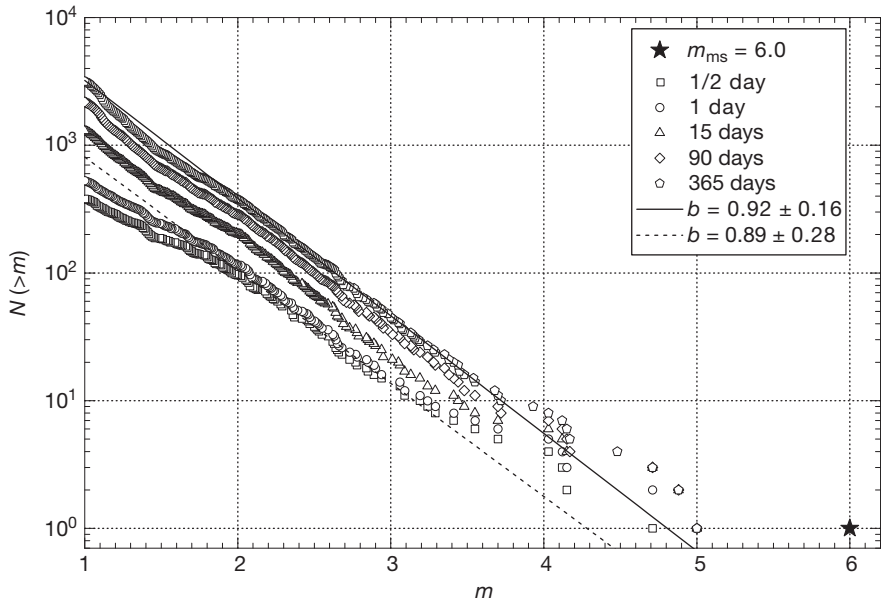


Figure 2 The frequency–magnitude distribution of aftershocks for several time intervals elapsed since the Parkfield mainshock. The Gutenberg–Richter scaling relation, eqn [1], is fit to obtain the b-values for 1/2 day ($b = 0.89 \pm 0.28$) and 365 days ($b = 0.92 \pm 0.16$) after the mainshock.

2002, 2003; Helmstetter and Sornette, 2003a; Kisslinger and Jones, 1991; Tsapanos, 1990; Vere-Jones, 1969).

A modified version of Båth's law has been proposed by Shcherbakov and Turcotte (2004) and is based on an extrapolation of the GR statistics for aftershocks. The magnitude of the 'largest' aftershock consistent with GR scaling for aftershocks is obtained by formally setting $N(\geq m)=1$ in eqn [1] with the result

$$a = bm^* \quad [3]$$

where m^* is the inferred magnitude of the 'largest' aftershock for the given aftershock sequence. In general, this extrapolated value will differ from the mean value of the largest aftershock obtained by averaging over an ensemble of mainshock–aftershock sequences having the same mainshock magnitude. If Båth's law is applicable to the inferred values of m^*, we can write

$$\Delta m^* = m_{\mathrm{ms}} - m^* \quad [4]$$

where Δm^* is approximately a constant. Substitution of eqns [3] and [4] into eqn [1] gives

$$\log_{10} N(\geq m) = b(m^* - m) \quad [5]$$

With b, m_{ms}, and m^* specified, the frequency–magnitude distribution of aftershocks can be determined using eqn [5].

The applicability of both forms of Båth's law is illustrated in **Figure 2** (Shcherbakov et al., 2006a). For the 2004 Parkfield earthquake, we have $m_{\mathrm{ms}}=6.0$ and $m_{\mathrm{as}}^{\mathrm{max}}=5.0$, from the fit of the GR relation [1] to the data we obtained $a=4.43\pm0.39$. From eqn [3], we obtain $m^*=4.8$. For this earthquake, the magnitudes of the largest actual aftershock and largest inferred aftershock are within 0.2 magnitude units. Shcherbakov and Turcotte (2004) studied ten large earthquakes that occurred in California between 1987 and 2003 with magnitudes equal to or greater than $m_{\mathrm{ms}}>5.5$. The mean difference in magnitudes between these mainshocks and their largest detected aftershocks was 1.16 with a standard deviation $\sigma=0.46$ (Båth's law). The mean difference in magnitudes between these mainshocks and their inferred 'largest' aftershocks was 1.11 with $\sigma=0.29$ (modified Båth's law). The modified version of Båth's law is a better approximation for these earthquakes than the original one.

Shcherbakov and Turcotte (2004) explained the general applicability of the modified form of Båth's law in terms of energy partitioning. The applicability of the law implies that the average ratio of the total energy radiated in an aftershock sequence to the energy radiated by the mainshock E_{as} is approximately a constant given by

$$\frac{E_{\mathrm{as}}}{E_{\mathrm{ms}}} = \frac{2b}{3-2b} 10^{-((3/2)\Delta m^*)} \quad [6]$$

Taking $\Delta m^*=1.2$ and $b=0.92$ for the Parkfield earthquake, we find from eqn [6] that $E_{\mathrm{as}}/E_{\mathrm{ms}}=0.025$.

4.24.2.3 The Omori Law

The temporal decay of aftershock activity following a mainshock also obeys a universal scaling law to a good approximation. This is known as the modified Omori law and can be written in the following form (Scholz, 2002; Utsu et al., 1995):

$$r(t) \equiv \frac{\mathrm{d}N}{\mathrm{d}t} = \frac{1}{\tau(1+t/c)^p} \quad [7]$$

where $r(t)\equiv \mathrm{d}N/\mathrm{d}t$ is the rate of occurrence of aftershocks with magnitudes greater than m, t is the time that has elapsed since the mainshock, τ and c are characteristic times, and the exponent p has a value near unity. The total number of aftershocks with magnitudes greater than m, $N_{\mathrm{tot}}(\geq m)$, in the time interval $[0, T]$ is obtained by integrating eqn [7] with the result (Shcherbakov et al., 2004)

$$N_{\mathrm{tot}}(\geq m) = \int_0^T \frac{\mathrm{d}t}{\tau(1+t/c)^p} = \begin{cases} \dfrac{c}{\tau(p-1)}\left[1-(1+T/c)^{1-p}\right], & \text{if } p\neq 1 \\ \dfrac{c}{\tau}\ln(1+T/c), & \text{if } p=1 \end{cases} \quad [8]$$

For most real aftershocks sequences, the exponent $p\neq1$; however, it can be pretty close to 1. Substitution of eqn [8] for $p\neq1$ into eqn [5] gives

$$c\left[1-(1+T/c)^{1-p}\right] = \tau(p-1)10^{b(m^*-m)} \quad [9]$$

This is an important relationship between the two characteristic times c and τ that appear in the modified Omori law, eqn [7].

In general, both c and τ can be functions of the lower cutoff magnitude m, that is, $c=c(m)$ and $\tau=\tau(m)$. To preserve the GR scaling, we assume

$$c(m) = c^* 10^{\gamma(m^*-m)} \quad [10]$$

$$\tau(m) = \tau^* 10^{-\alpha(m^*-m)} \quad [11]$$

where c^* and τ^* are characteristic timescales that correspond to $c^*=c(m^*)$ and $\tau^*=\tau(m^*)$, respectively, and γ and α are exponents. Substitution of eqns [10] and [11] into eqn [9] with the assumption that $T\gg c$ gives the requirements for $p>1$ that

$$c^* = (p-1)\tau^* \quad [12]$$

$$\gamma + \alpha = b \quad [13]$$

For $p<1$, one can similarly obtain

$$(c^*)^p T^{1-p} = (1-p)\tau^* \quad [14]$$

$$p\gamma + \alpha = b \quad [15]$$

We next consider the rates of aftershock occurrence at times $t\gg c$. In this limit, eqn [7] becomes

$$r(t\gg c, m) = \frac{1}{\tau(m)}\left[\frac{c(m)}{t}\right]^p \quad [16]$$

We further notice that the ratio of these asymptotic rates above magnitudes m and $m+dm$ satisfies a scaling relation given by

$$\frac{r(t\gg c, m)}{r(t\gg c, m+\mathrm{d}m)} = 10^{\beta\,\mathrm{d}m} \quad [17]$$

The quantity β is a b-value for aftershocks at large times. Substitution of eqns [10], [11], and [16] into eqn [17] gives the requirements that

$$p\gamma + \alpha = \beta \quad [18]$$

Combining eqns [13] and [18], one obtains for $p>1$

$$\alpha = \frac{pb - \beta}{p - 1} \quad [19]$$

$$\gamma = \frac{\beta - b}{p - 1} \quad [20]$$

In general, the b-value for all aftershocks is not equal to the b-value for aftershocks at times $t \gg c$ after the mainshock, that is, $\beta \neq b$.

The case $p < 1$ has to be considered separately. Comparing eqns [15] and [18], one can deduce for $p < 1$ that the two exponents are equal $\beta = b$. Relationships between scaling exponents such as that given in eqns [13], [15], [19], and [20] are often found near critical points defining the continuum phase transitions (Ma, 1976; Stauffer and Aharony, 1992).

The exponent β is the same exponent present in the productivity law analyzed by Felzer et al. (2004) and Helmstetter et al. (2005). This can be shown as follows (Bhattacharya et al., 2011):

$$K = \frac{c^p}{\tau} = \frac{(c^*)^p 10^{p\gamma(m^* - m)}}{\tau^* 10^{-\alpha(m^* - m)}} = \frac{(c^*)^p}{\tau^*} 10^{(p\gamma + \alpha)(m^* - m)} \quad [21]$$

where K is a productivity constant. Using eqn [18], we finally obtain

$$K \propto 10^{\beta(m_{ms} - \Delta m^* - m)} \quad [22]$$

This indicates that the productivity exponent is in fact the b-value of the sequence estimated from the frequency–magnitude statistics of aftershocks where the first several days are excluded from the sequence.

There are two limiting cases for the behavior of the modified Omori law. The first is to assume that the characteristic time c is a constant c_0, independent of the lower magnitude cutoff m of aftershocks that are considered.

In this case, from eqn [10], we have $\gamma = 0$, and from eqns [13] or [15] and [20], we obtain

$$b = \beta = \alpha \quad [23]$$

In this case, the two b-values, b and β, are equal. In this limit and also for $p > 1$ and $T \gg c$, the substitution of eqns [11] and [12] into eqn [7] gives

$$\frac{dN}{dt} = \frac{(p-1)10^{b(m^* - m)}}{c_0[1 + t/c_0]^p} \quad [24]$$

Similar results have been derived by Reasenberg and Jones (1989) and by Yamanaka and Shimazaki (1990).

In the second case, it is assumed that the characteristic time τ is a constant $\tau = \tau_0$, independent of the lower magnitude cutoff m. In this case, from eqn [11], we have $\alpha = 0$, and from eqns [13] and [19], we obtain $\gamma = b$ and

$$\beta = pb \quad [25]$$

As a representative example, we consider the rate of occurrence of aftershocks following the 2004 $m_{ms} = 6.0$ Parkfield earthquake (Shcherbakov et al., 2006a). The number of aftershocks per unit time $r \equiv dN/dt$ is given in **Figure 3** as a function of the time t after the mainshocks for several aftershock magnitude cutoffs, $m \geq 1.0$, 1.5, 2.0, 2.5, and 3.0. In **Figure 3**, the modified Omori law, eqn [7], is fit to the data for magnitudes above $m \geq 2.5$ using the maximum likelihood method (Guo and Ogata, 1997). This gives the estimates of the parameters of the modified Omori law: $p = 0.85 \pm 0.07$, $c(m \geq 2.5) = 0.002 \pm 0.009$ days, and $\tau = 0.001 \pm 0.002$ days. Assuming that both c and τ are functions of m as given by eqns [10] and [11] with $b = 0.92 \pm 0.16$, $m^* = 4.8$, and $p = 0.85 \pm 0.07$, it is possible to estimate $\gamma = 0.59$, $\alpha = 0.42$, and $c^* = 8.85$ s. The corresponding values of $c(m)$ and $\tau(m)$ are given in **Figure 3**.

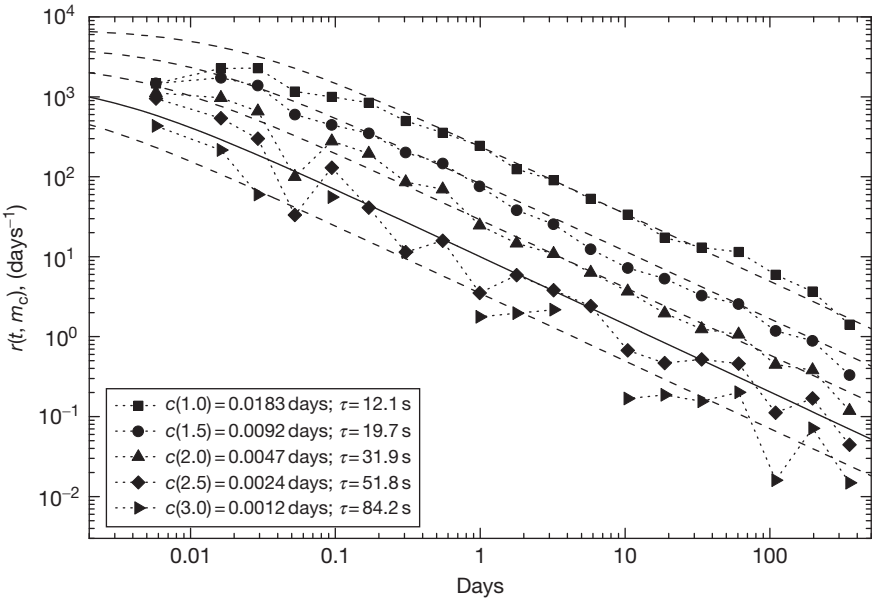


Figure 3 The rates of occurrence of aftershocks of the 2004 Parkfield earthquake with magnitudes greater than m_c in number per day, $r(t, m_c)$, are given for 365 days after the mainshock. Lower magnitude cutoffs were taken to be $m_c = 1.0$, 1.5, 2.0, 2.5, and 3.0. The value of $c(m_c = 2.5)$ has been obtained from the best fit. The rest of c's were calculated recursively.

It is seen that $c(m)$ is a strong function of m, whereas $\tau(m)$ has only a weak dependence on m. From eqn [18], we obtain $\beta = 0.92$; this compares with $b = 0.97 \pm 0.2$ estimated 1 day after the mainshock for the same interval of 365 days. Thus, β, the b-value for aftershocks at large times, is within error limits compared with the b-value for the whole sequence.

The fit of the modified Omori law to the data is quite good, and a large range of values of $c(m)$ are required. It is not possible to fit the data with a constant value of c. We note here that the catalog is not complete in the vicinity of the mainshock especially for small events. However, our form of the modified Omori law allowed us to recover the decay rate for small times after the mainshock. Some authors have argued that the entire rollover of aftershock rates at small times is an artifact of underrecording of early aftershocks (Kagan, 2004). However, our fit of the modified Omori law to the data can be accomplished using only the power-law decay curves.

Recent analyses of aftershock waveforms have shown that significant numbers of early events are missing in existing catalogs including the Parkfield aftershock sequence (Peng et al., 2006) and in Japan (Peng et al., 2007). This is generally ascribed to the lower sensitivity of recording instruments to smaller events in the wake of a large event and its large aftershocks. This masking of smaller magnitude aftershocks further complicates the problem of defining a minimum magnitude of completeness for the catalog. The analysis of four aftershock sequences in Japan generated by moderate mainshocks also revealed the existence of more aftershocks that were not reported in the JMA catalog (Enescu et al., 2009). By extracting the missing events, the authors established that the power-law decay started after several minutes after the mainshocks, suggesting the c-value was very small for those sequences. In another study, it was suggested that the stress regime was responsible for the observed variations in the c-values (Narteau et al., 2009). These studies reopen the question of the physical meaning of the c-value in the modified Omori law.

Results for several prominent earthquakes with well-documented aftershock sequences are given in **Table 1** (Nanjo et al., 2007; Shcherbakov and Turcotte, 2004; Shcherbakov et al., 2012, 2013b). In terms of magnitude, they range from the $m = 9.1$ Sumatra earthquake to the $m = 6.0$ Parkfield earthquake. The b-values range from 0.78 to 1.22. The values of Δm^* for the modified Båth's law range from 0.75 to 1.72. Values of the power-law exponent p in the modified Omori law range from 0.85 to 1.22.

Table 1 Parameters of the three empirical laws describing aftershock sequences of several prominent mainshocks

Mainshock	*Date*	m_{ms}	b	Δm*	p
Landers	26/06/1992	7.3	0.98	1.10	1.22
Northridge	17/01/1994	6.7	0.91	0.75	1.17
Kobe	17/01/1995	7.3	0.78	1.10	1.16
Hector Mine	16/10/1999	7.1	1.02	1.35	1.22
Tottori	06/10/2000	7.3	0.92	1.60	1.16
San Simeon	22/12/2003	6.5	1.00	1.10	1.12
Parkfield	28/09/2004	6.0	0.92	1.20	0.85
Sumatra	26/12/2004	9.1	1.13	1.72	1.10
Chile	27/02/2010	8.8	1.17	1.55	1.03
Darfield	04/09/2010	7.1	0.94	0.80	1.21
Tohoku	11/03/2011	9.0	1.22	1.56	1.02

We next consider the implications of the hypothesis that $\tau = \tau_0$ independent of the aftershock cutoff magnitude m. One implication of this hypothesis is that the GR frequency–magnitude scaling of aftershocks is not established instantaneously at the time of the mainshock. At short times after the mainshock with $t/c \ll 1$, eqn [7] reduces to

$$r \equiv \frac{dN}{dt} = \frac{1}{\tau_0} \quad [26]$$

Thus, τ_0 is the mean time delay before the first aftershock; also, there are very few small aftershocks at early times. At long times after the mainshock with $t/c \gg 1$, from eqn [7], we obtain

$$r \equiv \frac{dN}{dt} = \frac{1}{\tau_0}\left[\frac{c(m)}{t}\right]^p \quad [27]$$

This is the power-law decay of aftershock activity. With $\tau = \tau_0$, the dependence of $c(m)$ on m is obtained from eqn [9] with the result

$$c(m) = \tau_0(p-1)\left(\frac{10^{m^*}}{10^m}\right)^b \quad [28]$$

The equilibration time t_e for the establishment of GR distribution of aftershock magnitudes can be approximated by the time $t = t_e$ when the aftershock decay rates from eqns [26] and [27] are equal. The result is

$$t_e = c(m) \quad [29]$$

The equilibration time t_e corresponds to the characteristic time $c(m)$. Since $c(m)$ is a function of the cutoff magnitude m, so is the equilibration time $t_e(m)$.

We see from **Figure 3** that $c(m)$ systematically increases with decreasing lower cutoff magnitudes m. Thus, the equilibration time t_e also increases for the smaller magnitude aftershock cutoffs. The GR distribution is first established for the larger aftershocks and is subsequently established for smaller aftershocks.

The moment M of an aftershock can be related to its magnitude m by

$$\log_{10} M = d_1 m + d_2 \quad [30]$$

where d_1 and d_2 are constants. Kanamori and Anderson (1975) had established a theoretical basis for taking $d_1 = 3/2$. Substitution of eqn [30] into eqn [28] gives

$$c(m) = \tau_0(p-1)\left(\frac{M^*}{M}\right)^{2b/3} \quad [31]$$

where M^* is the moment of the 'inferred' largest aftershock with magnitude m^* and M is the moment of the cutoff magnitude m earthquake. Kanamori and Anderson (1975) had also shown that it is a good approximation to relate the moment M of an earthquake to the area A_r of the earthquake rupture by

$$M = \eta A^{3/2} \quad [32]$$

where η is a constant. This result is valid for earthquakes with rupture dimensions less than the seismogenic depth. Substitution of eqn [32] into eqn [31] with eqn [30] gives

$$t_e(A_r) = c(A_r) = \tau_0(p-1)\left(\frac{A_r^*}{A_r}\right)^b \quad [33]$$

where A_r^* is the rupture area of the 'inferred' largest aftershock with magnitude m^* and A_r is the rupture area of the cutoff magnitude m earthquake.

Setting $A_r = A_r^*$ in eqn [33], we have $t_e(m^*) = (p-1)\tau_0$. Since p is generally close to but greater than one, the equilibration time of the largest aftershock is shorter than the time delay τ_0 until the largest aftershock occurs. Thus, there is a time delay t_e for the establishment of the GR distribution only for small aftershocks.

We have shown that the time delays associated with aftershocks are well approximated by the modified Omori law, but why do the delays occur? When an earthquake occurs, there are regions where the stress is increased. This increase in stress is the fundamental cause of aftershocks. However, the systematic time delay before the occurrence of aftershocks requires an explanation. Das and Scholz (1981), Main (2000), and Ojala et al. (2003) had attributed this delay to stress corrosion combined with a critical stress intensity factor. Shaw (1993) had utilized a phenomenological approach to the dynamics of subcritical crack growth. A time delay is implicit in the empirically derived rate and state friction law. Dieterich (1994) had related the power-law decrease in aftershock activity to this law. There appear to be fundamental similarities between aftershock delays and the nucleation of bubbles in a superheated liquid. These similarities led Rundle (1989b), Rundle and Klein (1993), and Rundle et al. (1999) to relate aftershock sequences to the power-law scaling in the vicinity of a spinodal line. This association is also supported by the relationship between the three-dimensional (3D) spatial distributions of aftershocks and the 'backbone' of a 3D percolation cluster given by Robertson et al. (1995).

4.24.2.4 Accelerated Moment Release

There is accumulating evidence that there may be an increase in the number of intermediate-sized earthquakes prior to a large earthquake. The occurrence of a relatively large number of intermediate-sized earthquakes in northern California prior to the 1906 San Francisco earthquake was noted by Sykes and Jaumé (1990). Other studies have been carried out by Knopoff (1996) and by Jaumé and Sykes (1999).

It has also been observed that an increase in the seismic activity prior to large earthquakes takes the form of a power law. This was first proposed by Bufe and Varnes (1993). They considered the cumulative Benioff strain in a region defined as

$$\epsilon_B(t) = \sum_{i=1}^{N(t)} \sqrt{e_i} \quad [34]$$

where e_i is the seismic energy release in the ith precursory earthquake and $N(t)$ is the number of precursory earthquakes considered up to time t. It should be emphasized that the Benioff strain defined to be the square root of the seismic energy release is not an actual strain.

The precursory increase in seismicity is referred to as AMR. In terms of the cumulative Benioff strain, it is quantified as

$$\epsilon_B(t) = \epsilon_B(t_f) - B\left(1 - \frac{t}{t_f}\right)^s \quad [35]$$

where $\varepsilon_B(t_f)$ is the final cumulative Benioff strain at the time t_f when the large earthquake occurs, t is time measured forward from the beginning of AMR, B is a constant, and s is the exponent. Again, it should be emphasized that the cumulative Benioff strain is not proportional to the cumulative moment release since the earthquake moment is proportional to the earthquake energy release.

Examples of AMR have been given by Bowman et al. (1998), Bufe et al. (1994), Varnes and Bufe (1996), Brehm and Braile (1998, 1999a,b), Robinson (2000), Zöller et al. (2001), Main (1999), Bowman and King (2001), Yang et al. (2001), King and Bowman (2003), Bowman and Sammis (2004), Sammis et al. (2004), Jiang and Wu (2005, 2006), and Mignan et al. (2006), among others. Extensive reviews of AMR have been given by Mignan (2008, 2011).

Rundle et al. (2000) found that the distribution of values of the power-law exponent s for 12 earthquakes was $s = 0.26 \pm 0.15$. A systematic study of the optimal spatial region and magnitude range to obtain the power-law seismic activation has been carried out by Bowman et al. (1998). Four examples of their results are given in **Figure 4**, where the cumulative Benioff strain $\epsilon_B(t)$ has been correlated (solid line) with eqn [35]. Clear increases in seismic activity prior to these earthquakes are illustrated.

In obtaining the data given in **Figure 4**, Bowman et al. (1998) investigated circular regions about each of the characteristic earthquakes they considered. They determined the optimal radius for activation. Sornette and Sammis (1995) associated this precursory activation with the approach to a critical phase transition. Saleur et al. (1996a,b) interpreted this behavior in terms of correlation lengths that increase prior to the characteristic earthquake. Thus, we will refer to this radius of the optimal activation region as an activation correlation length (ACL).

Values of ACL as obtained by several authors are given in **Figure 5** as a function of earthquake magnitude and the characteristic rupture length. Although there is considerable scatter, there is a clear increase in the ACL with increasing earthquake magnitude. Also included for comparison is the relation

$$\xi = 10A_r^{1/2} \quad [36]$$

Many of the ACLs are greater than ten times the characteristic rupture length.

Dobrovolsky et al. (1979) and Keilis-Borok and Kossobokov (1990) reported a similar scaling for the maximum distance between an earthquake and its precursors using pattern recognition techniques. Large ACLs are also suggested by the remotely triggered aftershocks following the Landers, California, earthquake (Hill et al., 1993).

Zöller et al. (2001) directly estimated the ACL from earthquake catalogs using single-link cluster analysis (Frohlich and Davis, 1990). They studied 11 earthquakes in California with $m \geq 6.5$ since 1952 and found systematic increases in the correlation lengths prior to most of these earthquakes. Shcherbakov et al. (2006b) had analyzed the behavior of the two-point correlation length associated with foreshocks and aftershocks of several mainshocks. Goltz and Bose (2002) had

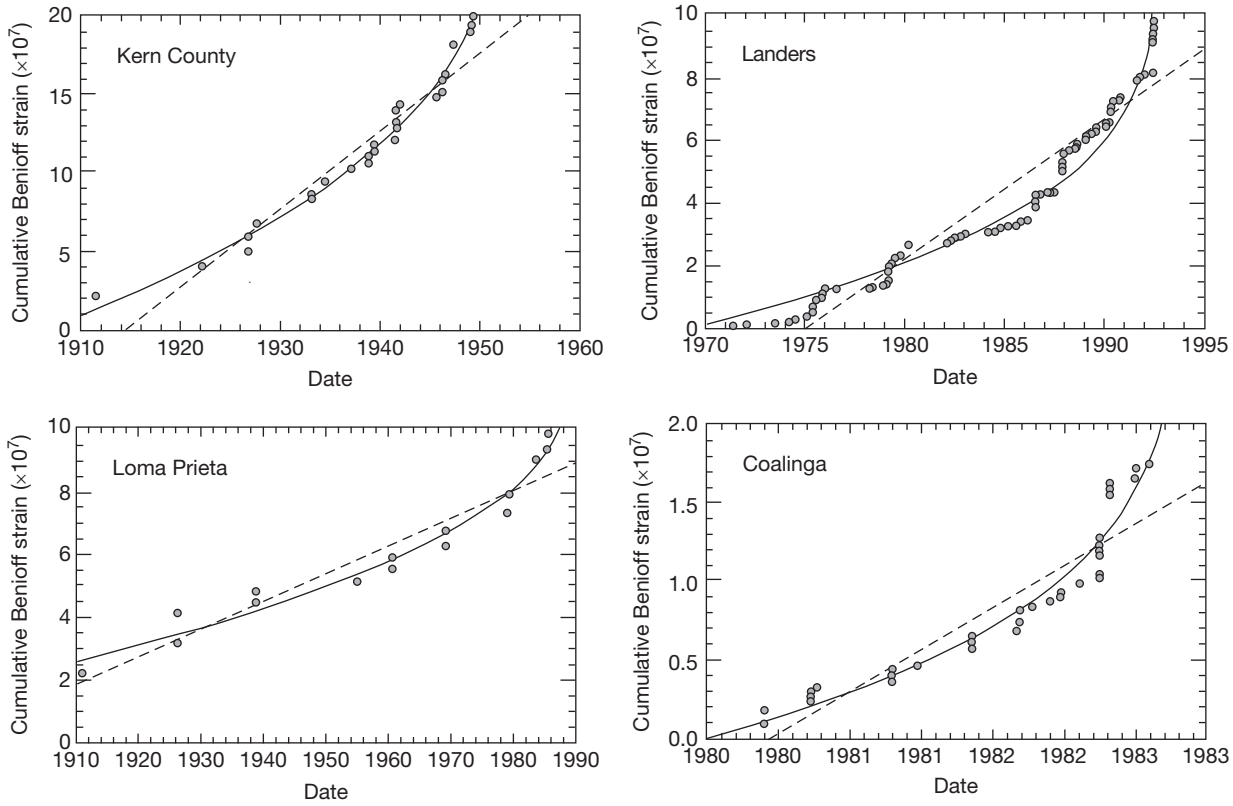


Figure 4 Data points are cumulative Benioff strains $\varepsilon_B(t)$ determined from eqn [34] prior to four major earthquakes in California (reproduced from Bowman DD, Ouillon G, Sammis CG, Sornette A, and Sornette D (1998)) An observational test of the critical earthquake concept. *Journal of Geophysical Research* 103(B10): 24359–24372.). Clear increases in seismic activity prior to the 1952 Kern County, 1989 Loma Prieta, 1992 Landers, and 1983 Coalinga earthquakes are illustrated. In each of the four examples, the data have been correlated (solid lines) with the power-law relation given in eqn [35]. The values of the power-law exponent s used in eqn [35] were 0.30 for Kern County, 0.18 for Landers, 0.28 for Loma Prieta, and 0.18 for Coalinga. Dashed straight lines represent a best-fit constant rate of seismicity.

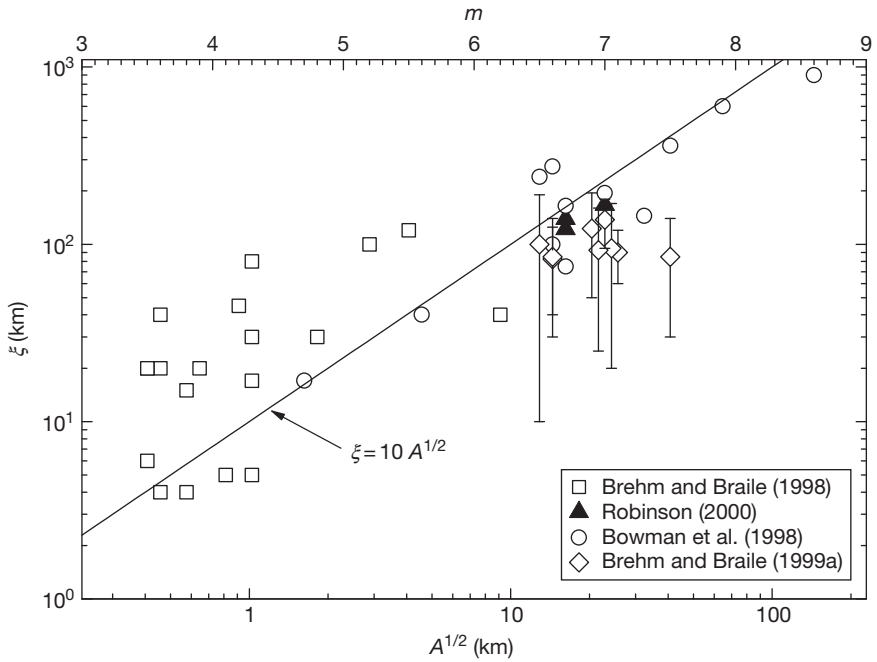


Figure 5 Correlation lengths ξ for precursory seismic activation are given as a function of the square root of the rupture area and the earthquake magnitude m. The open circles are the values given by Bowman et al. (1998), the open squares are values obtained for earthquakes in the New Madrid seismic zone by Brehm and Braile (1998), the error bars are limits obtained for earthquakes in the western United States by Brehm and Braile (1999a), and the solid triangles are values obtained for earthquakes in New Zealand by Robinson (2000). Reproduced from Rundle JB, Turcotte DL, Shcherbakov R, Klein W, and Sammis C (2003) Statistical physics approach to understanding the multiscale dynamics of earthquake fault systems. *Reviews of Geophysics* 41(4), 1019.

studied precursory seismic activation using configurational entropy.

Although there is considerable observational evidence for AMR, this phenomenon is certainly not universal for all earthquakes and systematic attempts to use AMR for earthquake predictions have not been successful. Critical evaluations of AMR have been given by Gross and Rundle (1998), Jaumé (2000), Vere-Jones et al. (2001), Zöller et al. (2001), Zöller and Hainzl (2001), Ben-Zion and Lyakhovsky (2002), Robinson et al. (2005), and Hardebeck et al. (2008). The concept of ACL is also controversial. Huc and Main (2003) and McKernon and Main (2005) had systematically looked for correlated seismicity in earthquake catalogs and had not found results consistent with ACL. There is also the question of the physical mechanisms associated with ACL. The lengths required are certainly greater than classic elastic length scales. However, Felzer and Brodsky (2006) had found that aftershocks systematically extend beyond elastic length scales. As a result, ACL may be associated with complex interactions between faults. A stress accumulation model for AMR has been proposed by Mignan et al. (2006, 2007).

4.24.2.5 Recurrence Times

A property of seismicity that is important in seismic hazard assessment is the statistical distribution of recurrence times of characteristic earthquakes. Examples of characteristic earthquakes include great earthquakes on the northern section of the San Andreas Fault (the 1906 San Francisco earthquake), the southern section of the San Andreas Fault (the great 1857 earthquake), and the Parkfield section of the San Andreas Fault.

Characteristic earthquakes are associated with quasi-periodicity but can also have considerable variability. A measure of the variability of recurrence times on a fault or fault segment is the coefficient of variation C_v (the ratio of the standard deviation σ to the mean μ). For strictly periodic earthquakes on a fault or fault segment, we would have $\sigma = C_v = 0$. For the random (i.e., exponential with no memory) distribution of recurrence times, we would have $C_v = 1$ (i.e., $\sigma = \mu$). Ellsworth et al. (1999) analyzed 37 series of recurrent earthquakes and suggested a provisional generic value of the coefficient of variation $C_v \approx 0.5$. A number of alternative distributions have been proposed for this purpose. These include the exponential (Poisson), log-normal, Brownian passage-time (inverse Gaussian), and Weibull distributions (Davis et al., 1989; Matthews et al., 2002; Ogata, 1999a; Sornette and Knopoff, 1997). We will primarily consider the Weibull distribution. The cumulative distribution function (cdf) for this distribution is given by

$$P(t) = 1 - \exp\left[-\left(\frac{t}{\tau}\right)^{\beta}\right] \qquad [37]$$

where $P(t)$ is the fraction of the recurrence times that are shorter than t and β and τ are fitting parameters. If $\beta = 1$, this is the Poisson (random) distribution.

Probably the best-studied sequence of characteristic earthquakes is the sequence that occurred on the Parkfield, California, section of the San Andreas Fault between 1857 and 2004 (Bakun et al., 2005). This is because the slip rate is relatively high (≈ 30 mm year^{-1}) and the earthquakes are relatively small ($m \approx 6.0$); thus, the recurrence times are relatively short (≈ 25 years).

Slip on the Parkfield section of the San Andreas Fault occurred during $m \approx 6$ earthquakes that occurred in 1857, 1881, 1901, 1922, 1934, 1966, and 2004. The cdf of Parkfield recurrence times ($t = 12$, 20, 21, 24, 32, 38 years) is given in Figure 6. The mean, standard deviation, and coefficient of variation of these recurrence times are $\mu = 24.5$ years, $\sigma = 9.25$ years, and $C_v = \sigma/\mu = 0.378$, respectively. Taking these values, the corresponding fitting parameters for the Weibull distribution are $\tau = 27.49$ years and $\beta = 2.88$. Using these values, the

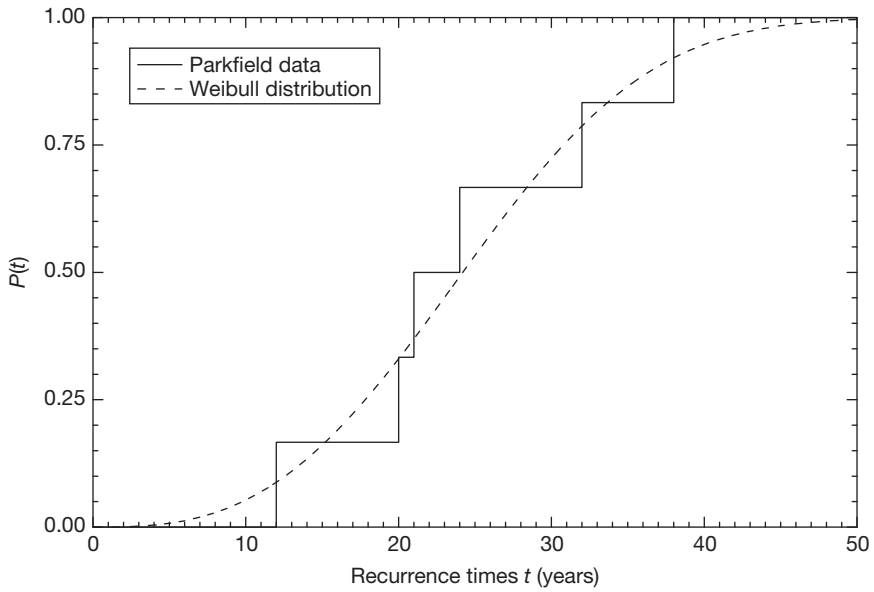


Figure 6 Cumulative distribution function of recurrence times of historical Parkfield mainshocks. The discontinuous line is the distribution of the six actual recurrence times. The continuous line is the best-fit Weibull distribution ($\mu = 24.5$ years, $\sigma = 9.25$ years, $C_v = 0.378$, $\tau = 27.5$ years, and $\beta = 2.88$). Reproduced from Abaimov SG, Turcotte DL, Shcherbakov R, et al. (2008) Earthquakes: Recurrence and interoccurrence times. *Pure and Applied Geophysics* 165(3–4): 777–795.

cdf from eqn [37] is also shown in Figure 6. Quite good agreement is observed.

Sequences of characteristic earthquakes that have been reliably dated are relatively short. In order to overcome this difficulty, Goltz et al. (2009) studied the recurrence time statistics of microrepeating earthquakes in California and Japan. They found good agreement with the Weibull distribution. Abaimov et al. (2007a) studied the recurrence time statistics of slip events on the creeping section of the San Andreas Fault in central California. Again, good agreement with the Weibull distribution was found.

Another important question is whether there is a scaling of seismicity that requires the validity of the Weibull distribution. The Weibull distribution is widely used in engineering to model the statistical distribution of failure times. Its applicability has been demonstrated by many actual tests (Meeker and Escobar, 1998).

It should be also pointed out that the stretched exponential distribution (a Weibull distribution with $0<\beta<1$) is found to be widely applicable in heterogeneous and homogeneous nucleation. In this context, its applicability is known as Avrami's law (Avrami, 1940). The similarities between nucleation problems (i.e., droplet formation in supercooled steam) and the nucleation of earthquakes have been discussed previously (Rundle et al., 2003). In Avrami's law, however, P is found to be less than one.

The sequence of recurrence times on a fault is a time series. Bunde et al. (2003, 2004), Altmann et al. (2004), Altmann and Kantz (2005), and Sanchez et al. (2002) had shown that time series that exhibit long-range correlations have recurrence time statistics that satisfy stretched exponential distributions. Studies have been carried out of the distribution of interval times between peaks over threshold for fractional Gaussian noises. These noises have a power-law dependence of the two-point correlation function C on the delay time s, $C(s)\approx s^{-\gamma}$. It was shown that the cumulative distribution of interval times satisfied the stretched exponential given by eqn [37] with $\beta=\gamma$. For fractional noises, $0<\gamma<1$ so that the distribution with $\beta<1$ has a fat tail. Bunde et al. (2003) argued against distributions in which the extreme event statistics become Poisson, that is, like the Brownian passage-time distribution.

A number of authors have related the spatial distribution of earthquakes to scale-free networks (Abe and Suzuki, 2003, 2004a,b; Baiesi, 2006; Baiesi and Paczuski, 2004, 2005; Console and Murru, 2001; Console et al., 2003b; Zaliapin et al., 2008).

4.24.2.6 Interoccurrence Times

Another property of seismicity that has been shown to satisfy a universal scaling law is the statistical distribution of interoccurrence times. All earthquakes in a specified region and specified time window T with magnitudes greater than m are considered to be point events. Based on studies of properties of California seismicity, a unified scaling law for the temporal distribution of earthquakes was proposed by Bak et al. (2002). The distribution of interoccurrence times between successive earthquakes was obtained by using as scaling parameters both a grid size over which the region was subdivided and a lower magnitude cutoff. Two distinct scaling regimes were found. For short times, aftershocks dominate the scaling properties of the distribution, decaying according to the modified Omori law. For long times, an exponential scaling regime was found that can be associated with the occurrence of mainshocks. To take into account the spatial heterogeneity of seismic activity, it has been argued that the second regime is not an exponential but another power law (Corral, 2003). An alternative approach to describe a unified scaling of earthquake recurrence times was suggested by Corral (2004a,b, 2005a,b, 2006a,b, 2007), where the distributions computed for different spatial areas and magnitude ranges were rescaled with the rate of seismic activity for each area considered. An analysis of the change in behavior between these two regimes based on a nonstationary Poisson sequence of events was carried out by Lindman et al. (2005). Related studies have been carried out by Carbone et al. (2005), Davidsen and Goltz (2004), and Livina et al. (2005a,b).

We illustrate this scaling law using the interoccurrence time distribution of Parkfield aftershocks. We define interoccurrence times between successive aftershocks as $\Delta t_i=t_i-t_{i-1}$, $i=1, 2,\ldots$ and study their statistical distribution over a finite time interval T; we use a lower magnitude cutoff m as a scaling parameter and study sequences with different m's. The normalized distribution (histogram) of interoccurrence times is given in Figure 7(a). The distributions with lower magnitude cutoffs have a shorter power-law scaling regime and start to roll over more quickly. This can be explained by the presence of more events in the sequences with lower magnitude cutoffs and as a result the shortening of the mean time intervals between events. Another scaling parameter that affects the rollover is the time interval T following the mainshock. An increase in T leads to the occurrence of longer time intervals Δt between events. In our analysis of aftershock interoccurrence times, we have used a time interval of $T=90$ days to try to eliminate the contribution from the background seismicity, which starts to influence the sequence in later times.

Shcherbakov et al. (2005b) showed that the observed distribution of interoccurrence times can be well approximated by a distribution derived from assuming that aftershocks follow nonhomogeneous Poisson statistics with the rate given by the modified Omori law [7] over a finite time interval T:

$$P_T(\Delta t)=\frac{1}{N}\left[\int_0^{T-\Delta t} r(s)r(s+\Delta t)\mathrm{e}^{-\int_s^{s+\Delta t} r(u)\mathrm{d}u}\mathrm{d}s+r(\Delta t)\mathrm{e}^{-\int_0^{\Delta t} r(u)\mathrm{d}u}\right] \quad [38]$$

where $N=\int_0^T r(u)\mathrm{d}u$ is the total number of events during a time period T. The plots of this equation for different magnitude cutoffs m are also given in Figure 7(a) as solid curves. Davidsen and Paczuski (2005) had also found a scaling for the distances between successive earthquake distributions in a region.

All the distributions at large interoccurrence times ($\Delta t>50$ s) collapse into a single functional form according to the following scaling law (Shcherbakov et al., 2005b):

$$P_{\tau,c,T}(\Delta t)=\frac{1}{\tau}\left(\frac{\Delta t}{\tau}\right)^{-\nu} f\left[\frac{\Delta t}{\tau}\left(\frac{T}{c}\right)^{-\delta}\right] \quad [39]$$

independent of the magnitude cutoff or time interval taken. Here, $f(x)$ can be approximated by the generalized gamma

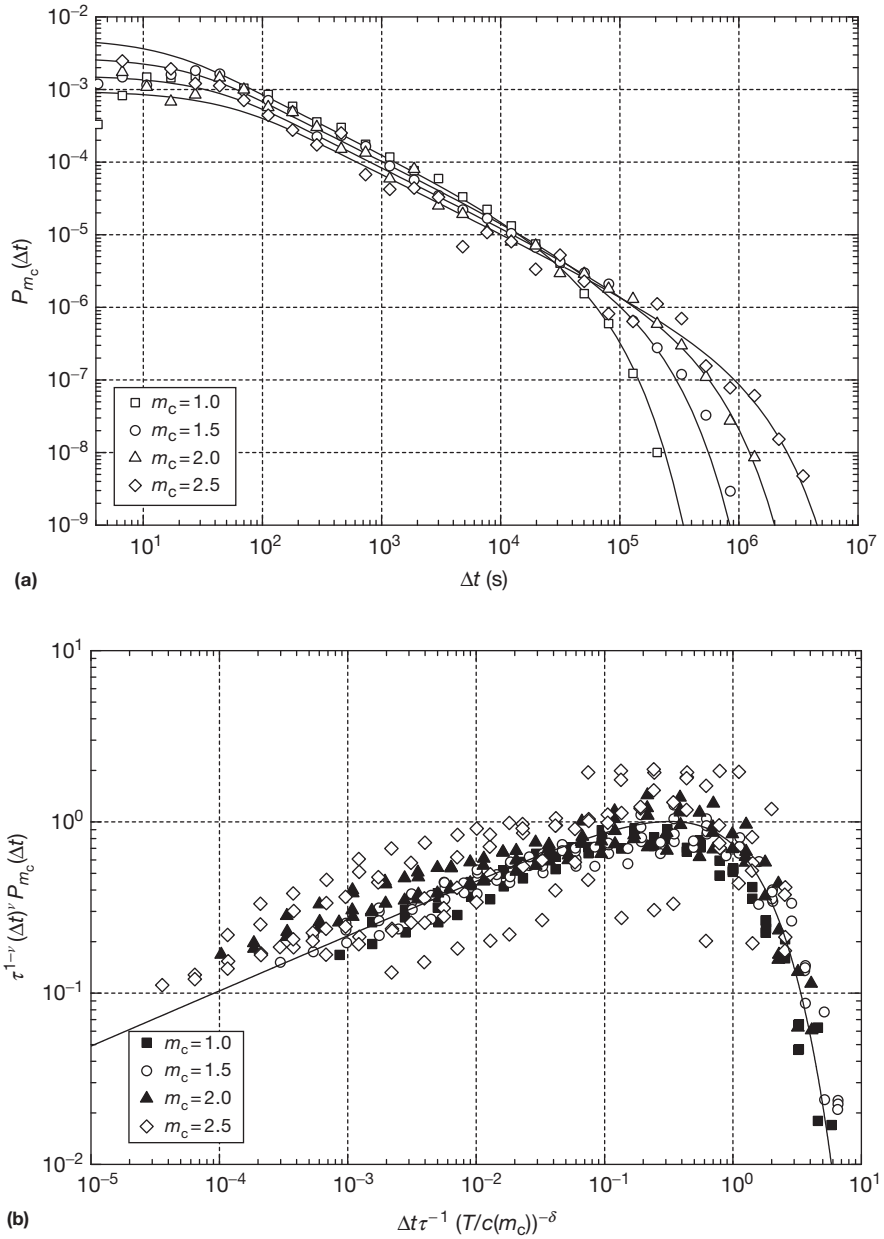


Figure 7 (a) Interoccurrence time distributions for the 2004 Parkfield aftershock sequence for different magnitude cutoffs $m_c = 1.0$, 1.5, 2.0, and 2.5. A time interval $T = 365$ days following the mainshock has been used. The solid lines have been computed using eqn [38]. (b) The scaling of interoccurrence time distribution for large times ($\Delta t > 50$ s) according to eqn [39]. Magnitude cutoffs $m_c = 1.0$, 1.5, 2.0, and 2.5 for the time intervals of 30, 60, 90, 180, 365, and 730 days are considered. The best data collapse is found for $\delta = 1.14$ and $\nu = 0.846$. The solid curve is an ft of the generalized gamma function, $f(x) = Ax^{\eta}\exp(-x/B)$, with $A = 2.0$, $B = 1.0$, and $\eta = 0.322$.

function, $f(x) = Ax^{\eta}\exp(-x/B)$, with $A = 2.0$, $B = 1.0$, and $\eta = 0.322$. It should be noted that ν and δ are independent exponents. We find that the best collapse of data occurs for $\nu = 1.14$ and $\delta = 0.846$. As was predicted by Shcherbakov et al. (2005b), in this case, $\delta \approx p$. This is shown in **Figure 7(b)**, which clearly shows that aftershocks can be approximated as a non-homogeneous Poisson process over several magnitude scales and underlines the self-similarity of aftershock sequences in both the time and magnitude domains. The same scaling law, however, does not apply for small interoccurrence times ($\Delta t < 50$ s). If a scaling does exist for the smaller interoccurrence intervals, it is not of the same mathematical form as for the larger interoccurrence times.

To constrain the effects of correlation and self-organization in seismogenic zones, one can study record-breaking earthquakes. A record-breaking event is defined to be an event larger or smaller than all previous events. A typical example is the sequence of record-breaking temperatures (either highest or lowest) on a specified day of the year at a specified site. Van Aalsburg et al. (2010) had studied the statistics of

record-breaking earthquake magnitudes. They considered global earthquakes for the period 1977–2006 and found agreement with the theory for a random process. Yoder et al. (2010) extended these studies to the record-breaking intervals between earthquakes in a specified region with magnitudes greater than a specified value. Davidsen et al. (2006, 2008) had applied the concept of record-breaking to spatial networks of recurring earthquakes. Shcherbakov et al. (2013a) studied the record-breaking events in nonlinear-driven threshold systems.

4.24.3 Self-Organizing Complex Systems

4.24.3.1 Fractality

Mandelbrot (1967) introduced the concept of fractality in terms of the length of a rocky coast line. Fractality implies power-law scaling and self-similarity. The basic definition of a fractal set is given by Turcotte (1997):

$$N_i = \frac{C}{r_i^D} \qquad [40]$$

where N_i is the number of objects with a characteristic linear dimension r_i, C is a constant of proportionality, and D is the fractal dimension.

Aki (1981) showed that the empirical GR frequency–magnitude scaling given in eqn [1] can be interpreted as a fractal relation. He used the empirical relation between earthquake magnitude m and the earthquake rupture area A_r obtained from eqns [30] and [32] to write

$$\log\left(\alpha A_r^{3/2}\right) = \frac{3}{2}m + d \qquad [41]$$

where α and d are constants. Combining eqns [1] and [41] gives

$$N = \frac{C'}{A_r^b} \qquad [42]$$

Comparing eqns [40] and [42], we can conclude that

$$D = 2b \qquad [43]$$

It must be emphasized that N is the cumulative number of earthquakes with characteristic linear dimension $A_r^{1/2}$ greater than the value specified whereas N_i in eqn [40] is noncumulative. It should also be emphasized that the integral of eqn [42] from 0 to ∞ diverges and cutoffs must be specified. For any natural process such as earthquakes, upper and lower cutoffs are certainly appropriate; for example, the validity of eqn [42] is greater than the atomic scale and less than the planetary scale.

4.24.3.2 Deterministic Chaos

Lorenz (1963) discovered deterministic chaos. He derived three coupled total differential equations as an approximation for thermal convection in a fluid layer heated from below. He showed that solutions in a parameter range had an exponentially diverging sensitivity to initial conditions. This divergent behavior is a characteristic of chaotic behavior. The chaotic behavior of these equations has been taken as evidence that the behavior of the atmosphere and oceans is chaotic.

Burridge and Knopoff (1967) introduced the slider-block model as a simple model for the behavior of a fault. One can show that a pair of connected slider blocks can exhibit deterministic chaos (Huang and Turcotte, 1990). Consider the behavior of a pair of slider blocks: a constant-velocity driver plate drags the blocks over a surface at a mean velocity, v. The two blocks are coupled to each other by a spring with spring constant k_c and to a driver plate by springs with spring constant k_L. Other model parameters are the block masses m and the coefficients of static and dynamic friction and μ_s and μ_d. Solutions are governed by two parameters, the stiffness of the system k_c/k_L and the ratio of static to dynamic friction coefficients μ_s/μ_d. Classic chaotic behavior is obtained for a wide range of parameter values as long as the symmetry is broken.

Slider-block models are a simple analogy to the behavior of faults in the Earth's crust. An important question is whether the chaotic behavior of a pair of slider blocks implies that seismicity is also an example of chaotic behavior. This question will be addressed in the next section. If seismicity is chaotic, it implies that the prediction of earthquakes is not possible in a deterministic sense. Only a probabilistic approach to the occurrence of earthquakes will be possible (Geller, 1997; Geller et al., 1997).

4.24.3.3 Criticality and SOC

The concept of SOC (Jensen, 1998; Turcotte, 1997, 1999b) evolved from the sandpile model proposed by Bak et al. (1988). In this model, there is a square grid of boxes, and at each time step, a particle is dropped into a randomly selected box. When a box accumulates four particles, they are redistributed to the four adjacent boxes, or in the case of edge boxes, they are lost from the grid. Since only nearest-neighbor boxes are involved in the redistribution, this is a cellular automaton model. Redistributions can lead to further instabilities and avalanches of particles in which many particles may be lost from the edges of the grid. This model was called a sandpile model because of the resemblance to an actual sandpile. The randomly dropped particles in the model are analogous to the addition of sand particles to a sandpile. The model avalanches are analogous to sand avalanches down the sides of the sandpile. Extensive analytical and numerical studies of the sandpile model were carried out by Kadanoff et al. (1989), Dhar (1990), Priezzhev et al. (1996), Ivashkevich and Priezzhev (1998), and Dhar (1999). It was found that the noncumulative frequency–area distribution of avalanches has a power-law distribution with a slope near 5/4 (Priezzhev et al., 1996). A modification of the sandpile model that utilizes a fractal distribution of box sizes reproduces GR statistics for aftershocks (Barriere and Turcotte, 1991, 1994; Henderson et al., 1994; Huang et al., 1998). A model with two state variables was also introduced to illustrate the concept of SOC (Shcherbakov and Turcotte, 2000).

Another example of SOC is observed in the behavior of large arrays of interconnected slider blocks (Burridge and Knopoff, 1967). The two-block model was considered in the last section. The standard multiple slider-block model consists of a square array of slider blocks as illustrated in **Figure 8**. Each block with mass m is attached to the loader plate with a loader spring, spring constant k_L. Adjacent blocks are attached to each other with a connector spring, spring constant k_c. The constant-velocity driver plate is pulled forward with velocity v. A block remains stationary as long as the net force on the block is less than the static resisting

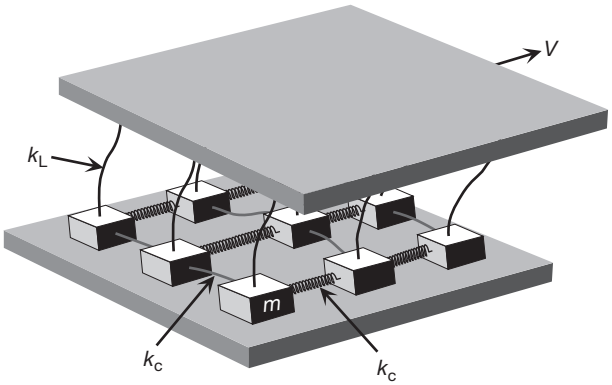


Figure 8 Illustration of the two-dimensional slider-block model. An array of blocks, each with mass m, is pulled across a surface by a driver plate at a constant velocity, V. Each block is coupled to the adjacent blocks with either leaf or coil springs (spring constant k_c) and to the driver plate with leaf springs (spring constant k_L). Reproduced from Malamud B.D. and Turcotte D.L. 1999. Self-Organized Criticality Applied to Natural Hazards. *Natural Hazards* v.20, p93–116

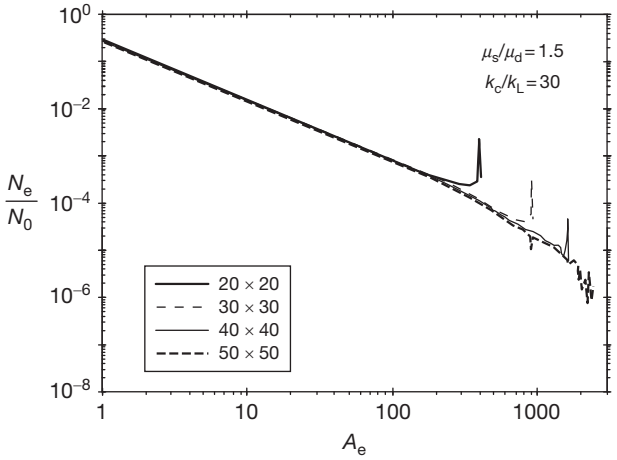


Figure 9 Results of simulations for a 2D slider-block model with multiple blocks (reproduced from Huang J and Turcotte DL (1992)) Chaotic seismic faulting with a mass-spring model and velocity-weakening friction. *Pure and Applied Geophysics* 138(4): 569–589. The ratio of the number of slip events, N_e, with area A_e, to the total number of slip events, N_0, is plotted against A_e – the number of blocks involved in an event. Results are given for systems with stiffness $k_c/k_L = 30$, ratio of coefficients of friction $\mu_s/\mu_d = 1.5$, and grid sizes 20×20, 30×30, 40×40, and 50×50. The peaks at $A_e = 400$, 900, and 1600 correspond to catastrophic slip events involving the entire system.

force. Again, two parameters determine the behavior of the system, the ratio of static and dynamic coefficients of friction, μ_s/μ_d, and the stiffness of the system, k_c/k_L. The slider-block model (Burridge and Knopoff, 1967) can be considered as a simple realization of a fault surface with threshold avalanche events representing the analogue of earthquakes.

Carlson and Langer (1989) considered long linear arrays of slider blocks with each block connected by springs to the two neighboring blocks and to a constant-velocity driver. They used a velocity-weakening friction law and considered up to 400 blocks. Slip events involving large numbers of blocks were observed, the motions of all blocks involved in a slip event were coupled, and the applicable equations of motion had to be solved simultaneously.

The system was completely deterministic. Frequency–area statistics were obtained for slip events. The events fell into two groups: Smaller events obeyed a power-law (fractal) relationship with a slope near unity, but there were an anomalously large number of large events that included all the slider blocks. The observed behavior was characteristic of SOC. The motion of the driver plate is the steady input. The slip events are the 'avalanches' with a power-law (fractal) frequency–size distribution.

Nakanishi (1990) studied multiple slider-block models using the cellular automaton approach. A linear array of slider blocks were considered, but only one block was allowed to move in a slip event. The slip of one block could lead to the instability of either or both of the adjacent blocks, which would then be allowed to slip in a subsequent step or steps, until all blocks were again stable. Brown et al. (1991) proposed a modification of this model involving a 2D array of blocks. The use of this approach greatly reduces the complexity of the calculations, and the results using the two approaches are generally very similar. A wide variety of slider-block models have been proposed and studied; these have been reviewed by Rundle et al. (2003).

Huang and Turcotte (1992) carried out a large number of simulations on a square array of blocks using stick–slip friction and the cellular automaton approach. Their noncumulative frequency–area statistics for model slip events are given in Figure 9. The number of slip events per time step N_e/N_T with area A_e is given as a function of A_e. The simulation results are well fit by the power-law scaling relation (eqn [40]) with $D = 2.72$. For stiff systems (k_c/k_L large), the entire grid of slider blocks is strongly correlated and large slip events including all the blocks occur regularly. These are the peaks for $A_e = 400$, 900, and 1600 blocks, illustrated in Figure 9. For soft systems (k_c/k_L relatively small), no large events occur.

Abaimov et al. (2007b, 2008, 2009) utilized a stiff slider-block model to study the recurrence time statistics of characteristic earthquakes. Slip events that included all slider blocks modeled characteristic earthquakes. Small slip events satisfying power-law frequency–area statistics modeled the GR scaling for background seismicity. The recurrence time distribution of characteristic (system-wide) slip events was found to be in excellent agreement with the Weibull distribution with a coefficient of variation $C_v = 0.48$. This value is in good agreement with values found for characteristic earthquakes as discussed in Section 4.24.2.6.

A cellular automaton version of the slider-block model was introduced by Olami et al. (1992) to analyze the behavior of massless blocks interconnected by springs and pulled over a surface with friction. The Olami–Feder–Christensen (OFC) model is locally nonconservative, which results in different regimes of behavior depending on the degree of nonconservation specified by the parameter α.

In the model, each site (i, j) of a square lattice is assigned a continuous variable $h_{ij} \in [0, h_c]$, which defines a stress level. The model is driven uniformly with a constant rate until one of the sites reaches the threshold value $h_c = 1.0$. This site becomes unstable and redistributes stress to its nearest neighbors and is reset to zero. The amount of stress transferred to

each nearest-neighbor site is αh_{ij}, where $0 < \alpha \leq 0.25$ defines a degree of nonconservation. This transfer may cause some of the neighbors to topple, initiating an avalanche. The model is locally nonconservative for $\alpha < 0.25$. The parameter α plays an important role in the dynamics of the model with smaller values of α leading to longer transient regimes (Wissel and Drossel, 2006) and also affecting its critical properties. The question of criticality of the OFC model still remains open (Wissel and Drossel, 2006).

The model produces rich behaviors that can be explored in order to understand its dynamics. After passing the transient regime, the model forms highly correlated spatial structures with power-law distributed patches of different heights. The presence of the nonconservation parameter α introduces several scales over which the distributions of avalanches are different at different scales. Large patches are mostly responsible for large avalanches, while small avalanches are generated by an interplay between small patches (Wissel and Drossel, 2006). The effects of correlations and the structure of record-breaking avalanches in the OFC model were studied by Shcherbakov et al. (2013a).

Recent results concerning the spatiotemporal organization of avalanches in the OFC model show that the model generates, in addition to regular avalanches, near periodic avalanches associated with asperity structures formed as a highly concentrated stress state (Kawamura et al., 2010; Kotani et al., 2008). These asperity events form spatiotemporal clusters with pronounced characteristic features. The processes of synchronization and desynchronization in the OFC model with short-term correlations among events were also observed by Hergarten and Krenn (2011). It was suggested that synchronization in the model drives it into a state with scale-invariant power-law distributed events, while desynchronization results in the occurrence of aftershocks and foreshocks.

It was also suggested that the OFC model can exhibit the maximum in the entropy production rate for the range of the parameter $\alpha \in [0.15, 0.2]$. This was associated with finite spatial memory of past avalanches associated with asperity events (Main and Naylor, 2008, 2010; Naylor and Main, 2008). There is a tendency of similar large avalanches to reoccur as the model is driven forward. These avalanches rupture approximately the same area and occur quasiperiodically in time. The range of the parameter α where the maximum of the entropy production rate is achieved corresponds to the subcritical regime of the OFC model. A critical point was suggested at $\alpha = 0.25$ with a diverging correlation length and average avalanche sizes (Main and Naylor, 2010).

A simple inverse-cascade model can explain the behavior of slider-block arrays (Turcotte, 1999a,b). A metastable cluster is the region over which a slip event propagates once it has been initiated. At any one time, there can be many metastable clusters on the grid. The metastable regions grow by coalescence, which is self-similar and gives power-law scaling. Avalanches sample the distribution of smaller clusters and, at the same time, remove the largest clusters. Gabrielov et al. (1999) and Yakovlev et al. (2005) had shown that the coalescence of clusters forms an inverse cascade that is identical to the formation of fractal drainage networks. The primary mechanism for the coalescence of clusters is bridging of gaps between clusters. The inverse cascade is further quantified by introducing the concept of a cluster order, which is identical to the classification of stream networks. The joining of streams as stream networks evolve is analogous to the coalescence of clusters as they grow. Streams can be classified according to both the primary branching and side or secondary branching. Clusters can be similarly defined. When two clusters of equal order coalesce, it is analogous to the joining of two streams of the same order. When a cluster of lower order coalesce with a cluster of higher order, it is analogous to the side branching between a stream of lower order and a stream of higher order. In the slider-block model, slip events sample the population of smaller clusters but do not deplete their number. Metastable blocks cascade from smaller clusters to larger clusters until the cascade is terminated in the largest slip events. These largest slip events define the upper limit of the power-law frequency–area scaling of both metastable clusters and slip events. The cascade is self-similar and the populations of clusters in terms of primary and secondary coalescences are identical to the populations of branches in self-similar drainage networks.

4.24.4 Relevant Models

4.24.4.1 Cellular Automaton Models

Cellular automaton models involve a grid of sites. At each time step, the model is updated by a prescribed set of rules. Interactions among prescribed sites are limited to a neighborhood, usually nearest neighbors. The two models discussed in the previous section, the sandpile model and the OFC model, are examples of cellular automaton models. In this section, we will consider a third example, the forest-fire model (Drossel and Schwabl, 1992; Henley, 1993). The forest-fire model consists of a square grid of sites; at each time step, either an attempt is made to plant a tree on a randomly selected site or a model match is dropped on the site. A tree is planted if a site is not occupied by a tree. If a match is dropped on a site occupied by a tree, that tree and all adjacent trees are destroyed in a model fire. The sparking frequency is the inverse of the number of attempted tree drops before a model match is dropped. It is found that the frequency–area distribution of fires over a range of sizes is power law with a slope near unity. Thus, the forest-fire model behaves very much like the sandpile and the slider-block models. Like the sandpile model, sites are selected randomly so that it is a stochastic model. Like the slider-block model, there is a tuning parameter, the sparking frequency. If the sparking frequency is high, fires prevent the formation of large clusters. The power-law cascade of planted trees from small to large clusters is terminated with a maximum cluster size that is a fraction of the grid size. If the sparking frequency is low, large system size clusters occur.

Another model that is widely studied in statistical physics is the site-percolation model (Stauffer and Aharony, 1992). The site-percolation model can be regarded as the forest-fire model without fires. This is a transient problem as the number of trees on the grid increases monotonically until every site is occupied by a tree. However, there is a well-defined critical point during this evolution. This critical point occurs when a cluster of trees span the grid. Monte Carlo simulations with large square grids show that a spanning cluster is formed when 59.275% of the grid sites are occupied by trees. At this critical point, the

frequency–area distribution of tree clusters is power law with a slope of 2.055. This similarity to the forest-fire model only occurs at the critical point. At other times during the transient planting of trees, the distribution is not power law (Turcotte, 1999b). The tuning parameter in this critical-point problem is the fraction of sites occupied by the spanning cluster. The critical value of this tuning parameter for site percolation on a square grid is $p_c = 0.59275$.

In order to study the behavior of characteristic earthquakes, a model has been developed (Newman and Turcotte, 2002) that combines aspects of the forest-fire and site-percolation models. In this model, fires only burn spanning (percolating) clusters. Trees are planted on unoccupied sites until a cluster spans the grid; at this time step, the entire cluster is removed and tree planting continues. The inverse cascade of trees from smaller to larger clusters gives a power-law distribution of cluster sizes at all times.

The fires in this model are somewhat analogous to characteristic earthquakes on a major fault or fault segment. The planting of tree corresponds to the tectonic increase in stress. The variations in cluster structure correspond to stress variations in the crust. The variability of the sizes is analogous to the variability in characteristic earthquakes and the associated variability in the recurrence times between characteristic earthquakes. In this model, there is a memory of the structure of previous fires for any given fire; each characteristic earthquake is influenced by the structure of previous characteristic earthquakes. The cdf of recurrence intervals (in time steps) between fires is given in **Figure 10**. The mean number of time steps between fires is $\mu = 7385$, the standard deviation is $\sigma = 2908$, and the coefficient of variation is $C_v = 0.3938$. Also included is the cdf for the Weibull distribution, eqn [37], with the same μ and coefficient of variation C_v as the model results; this required $\tau = 8300$ and $\beta = 2.74$. It is also interesting to note that the $C_v = 0.394$ for the recurrence times of the model fires is very similar to the $C_v = 0.378$ for the Parkfield earthquakes.

The applicability of the Weibull distribution to the recurrence times for earthquakes was discussed in **Sections 4.24.2–4.24.2.5**. It was pointed out that the Weibull distribution is widely used for failure time statistics in engineering (Meeker and Escobar, 1998).

4.24.4.2 Fiber-Bundle Models

A standard model to simulate a failure problem in engineering is the fiber bundle or stranded cable. The generally accepted approach to the dynamic time-dependent failure of a fiber bundle is to specify an expression for the rate of failure of fibers (Coleman, 1958). The form of this breakdown rule is given by

$$\frac{dN}{dt} = -\nu_0 N \left(\frac{\sigma}{\sigma_0}\right)^p \qquad [44]$$

where N is the number of unbroken fibers at time t, σ is the applied stress that can be a function of time, ν_0 is the reference hazard rate at the reference stress σ_0, and p is the power-law exponent.

Since in the fiber-bundle model we assume that the stress σ on all fibers is equal, the fibers do not interact. Thus, it is appropriate to assume an ergodic hypothesis. Instead of applying our model to a fiber bundle with N_0 fibers, we apply it to a single fiber that is replaced N_0 times. The single fiber is analogous to a fault, and the failure of the fiber is analogous to an earthquake. The distribution of failure times of the fiber is analogous to the distribution of earthquake recurrence times on a fault. The cdf of failure times of fibers P is related to N by

$$P(t) = 1 - \frac{N(t)}{N_0} \qquad [45]$$

where N_0 is the total number of earthquakes. We further assume that after each fiber failure, the stress on the fiber is reset to zero and increases linearly with time

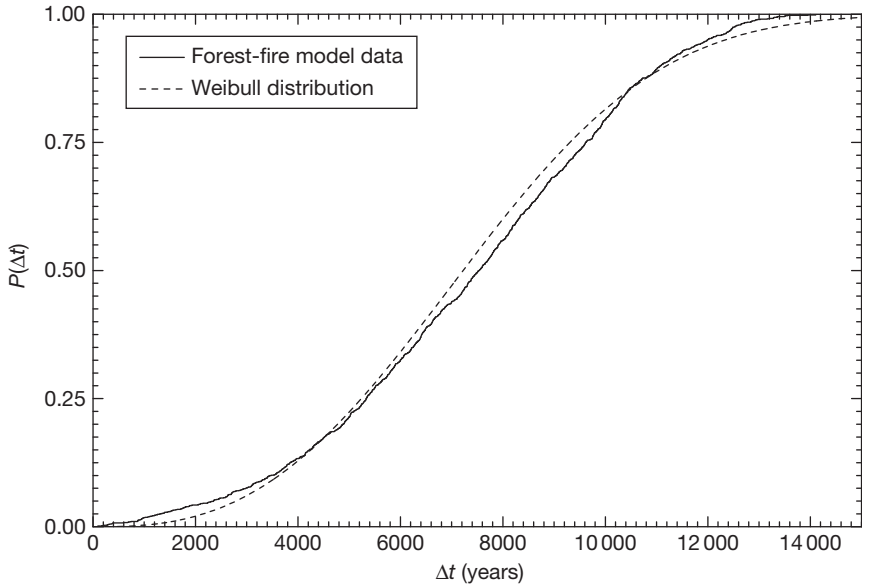


Figure 10 Cumulative distribution function of recurrence intervals (in time steps) between model fires. The dashed line is the best-fit stretched exponential distribution ($\mu = 7385$, $\sigma = 2908$, $C_v = 0.3938$, $\tau = 8300$, and $\beta = 2.74$).

$$\sigma(t) = \sigma_0 \frac{t}{t_0} \quad [46]$$

This represents the linear increase in the tectonic stress on a fault after an earthquake. In order to relate the constants v_0 and p in the fiber-bundle model to the constants β and τ in the Weibull distribution, we write

$$p = \beta - 1 \quad [47]$$

$$v_0 = \frac{\beta}{\tau} \quad [48]$$

Substitution of eqns [45] and [48] into eqn [44] gives

$$\frac{dP(t)}{dt} = \frac{\beta}{\tau}\left(\frac{t}{\tau}\right)^{\beta-1}[1 - P(t)] \quad [49]$$

Integrating with the initial condition that $P(0) = 0$ at $t = 0$ gives

$$P(t) = 1 - \exp\left[-\left(\frac{t}{\tau}\right)^{\beta}\right] \quad [50]$$

This is the cumulative form of the Weibull distribution given in eqn [37].

4.24.4.3 Numerical Simulations

Simulation-based approaches to forecasting and prediction of natural phenomena have been used with considerable success for weather and climate. Turbulent phenomena are represented by parameterizations of the fluid dynamics, and the equations are typically solved over spatial grids having length scales of tens to hundreds of kilometers. Even though the behavior of the atmosphere and oceans is recognized to be chaotic, numerical weather forecasts are generally quite accurate for periods of 48–72 h. A particular example is the probabilistic forecasts of hurricane tracks.

Ideally, similar simulations can be carried out to assess the earthquake hazard. However, tectonic models involving finite strain are not feasible at this time. Such models would require both block motions (displacements on faults) and continuum deformation (granulation). It is not possible to evolve block motions due to geometrical incompatibilities, and there will be either overlaps or gaps.

A more limited simulation model for distributed seismicity on the San Andreas and adjacent faults in southern California was given by Rundle (1988). This model included stress accumulation and release as well as stress interactions including the San Andreas and adjacent faults. The model was based on a set of mapped faults with estimated slip rates, a prescribed plate tectonic motion, earthquakes on all faults, and elastic interactions.

An updated version of the model has been developed (Rundle et al., 2001, 2002a, 2004, 2005, 2006). This version, called Virtual California, is a 3D earthquake simulation code characterized by the following:

1. A fault model is prescribed, and both strike-slip and dip-slip faults are included. A California fault model includes about 180 faults and fault segments. Each fault or fault segment is meshed into an array of elements, typically 3×3 km. In order to constrain the behavior of the faults, a mean slip velocity and a mean recurrence interval between earthquakes are prescribed for all faults and fault segments.
2. The Virtual California code embeds the prescribed faults in an elastic half-space. A fault displacement on a fault element produces stress changes on all other fault elements. These stress interactions are obtained using dislocation theory. Stress accumulation on a fault is by means of backslip. The displacement across a fault element is increased using the mean slip velocity for the element. A model earthquake occurs on a fault element when the shear stress difference across the fault exceeds the prescribed static coefficient of friction. When a fault element fails, the stress on the element is reduced to the dynamic coefficient of friction. Stress transfer from the failed element often induces the failure of adjacent elements. The number of elements that fail during the propagating rupture determines the magnitude of the resulting model earthquake. The prescribed static coefficient of friction is 'tuned' to give the prescribed recurrence interval between earthquakes.
3. The output of the simulation is a catalog of synthetic earthquakes on the prescribed fault system. This catalog can be used to generate a variety of earthquake statistics, that is, frequency–magnitude and recurrence and interoccurrence times.

Yakovlev et al. (2006) had used Virtual California simulations to study earthquake recurrence times on the San Andreas Fault system. The statistical distribution of recurrence times on the northern San Andreas Fault (site of the 1906 San Francisco earthquake) was obtained from a 1 000 000-year simulation. The cdf for 4606 simulated earthquakes with $M > 7.5$ on this section is given in **Figure 11**. The mean recurrence time is $\mu = 217$ years, the standard deviation is $\sigma = 114.7$ years, and the coefficient of variation is $C_v = 0.528$. Also included is the cdf for the Weibull distribution with the same mean μ and standard deviation σ as the simulation results. This required $\tau = 245$ years and $\beta = 1.97$. It is seen that the Weibull distribution is in quite good agreement with the results of the simulation.

Yikilmaz et al. (2010) utilized Virtual California simulations to study the influence of an asperity on earthquake ruptures. They divided a straight strike-slip fault into two long weak segments and a third strong central segment (the asperity). Limit cycle behavior was observed that produced earthquake sequences very similar to the historical sequence of 13 great earthquakes along the Nankai Trough, Japan. Yikilmaz et al. (2011) generated synthetic catalogs of seismicity in northern California using a composite simulation of earthquakes. A Virtual California simulation generated characteristic earthquakes on prescribed faults for a period of 100 000 years. These earthquakes were predominantly in the range $m = 6.0$–8.0. In order to model seismicity on unmapped faults, background seismicity was introduced randomly in time with GR scaling and a prescribed proximity to the mapped faults. Aftershocks were also introduced using the branching aftershock sequence (BASS) model (see **Section 4.24.4.5**).

Other simulation models of seismicity have been carried out. One example is the standard physical earth model, developed by Ward (1992) and applied to seismicity associated with subduction at the Middle America Trench. This model was

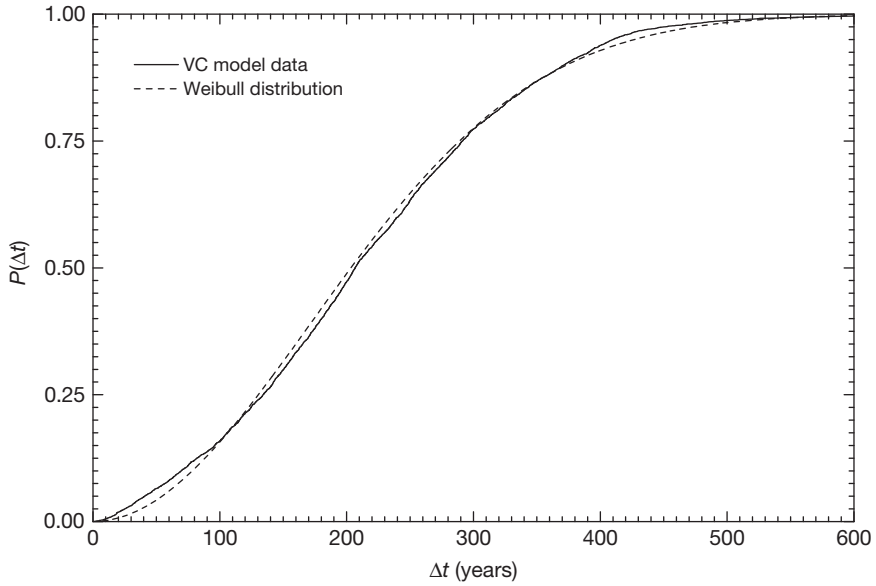


Figure 11 Cumulative distribution function of recurrence times for 4096 simulated earthquakes with $m > 7.5$ on the northern San Andreas Fault. The dashed line is the best-fit stretched exponential distribution ($\mu = 217$ years, $\sigma = 115$ years, $C_v = 0.528$, $\tau = 245$ years, and $\beta = 1.97$).

further developed and applied to the entire San Andreas system by Goes and Ward (1994), to the San Andreas system in southern California by Ward (1996), and to the San Andreas system in northern California by Ward (2000).

Other simulation results for California have been given by Dieterich and Richards-Dinger (2010) and Pollitz (2011). Although the four different models for seismicity in California have many significant differences, the statistical results are quite similar. Robinson and Benites (1995, 1996) had generated simulation results for New Zealand.

4.24.4.4 Damage Mechanics Models

One approach to the complex brittle deformation of a solid is to use damage mechanics. This approach was originally developed to quantify the irreversible brittle deformation of solids associated with microcracking. Several empirical continuum damage mechanics models were introduced and are used in civil and mechanical engineering (Kachanov, 1986; Krajcinovic, 1996; Shcherbakov and Turcotte, 2003; Yakovlev et al., 2010). Damage mechanics has also been applied to the brittle deformation of the Earth's crust by a number of authors (Ben-Zion and Lyakhovsky, 2002, 2006; Lyakhovsky and Ben-Zion, 2008; Lyakhovsky et al., 1997, 2001; Turcotte and Shcherbakov, 2006; Turcotte et al., 2003). In this application, the deformation and damage are associated with earthquakes.

The application of continuum damage mechanics can be illustrated by considering a rod in a state of uniaxial stress. For an elastic material, Hooke's law is applicable and is written in the form

$$\sigma = E_0 \epsilon \qquad [51]$$

where ε is a strain and E_0 is Young's modulus of the undamaged material.

In this chapter, we will consider a model of continuum damage mechanics as introduced by Shcherbakov et al. (2005a). If the stress is less than the yield stress, eqn [51] is assumed to be valid. If the stress is greater than the yield stress, $\sigma > \sigma_y$, a damage variable α is introduced according to

$$\sigma - \sigma_y = E_0(1-\alpha)(\epsilon - \epsilon_y) \qquad [52]$$

where $\sigma_y = E_0 \epsilon_y$. When $\alpha = 0$, eqn [52] reduces to eqn [51] and linear elasticity is applicable, as failure occurs. Increasing values of α in the range $0 \le \alpha < 1$ quantify the weakening (decreasing E) associated with the increase in the number and size of microcracks in the material. Several authors (Krajcinovic, 1996; Turcotte et al., 2003) have shown a direct correspondence between the damage variable α in a continuum material and the number of surviving fibers N in a fiber bundle that originally had N_0 fibers, $\alpha = 1 - N/N_0$.

To complete the formulation of the damage problem, it is necessary to specify the kinetic equation for the damage variable. In analogy to Lyakhovsky et al. (1997), we take

$$\frac{d\alpha(t)}{dt} = 0, \quad \text{if } 0 \le \sigma \le \sigma_y \qquad [53]$$

$$\frac{d\alpha(t)}{dt} = \frac{1}{t_d}\left[\frac{\sigma(t)}{\sigma_y} - 1\right]^{\rho}\left[\frac{\epsilon(t)}{\epsilon_y} - 1\right]^{2}, \quad \text{if } \sigma > \sigma_y \qquad [54]$$

where t_d is a characteristic timescale for damage and p is a constant to be determined from experiments. The power-law dependence of $d\alpha/dt$ on stress (and strain) given in the preceding text must be considered empirical in nature. However, a similar power dependence of dN/dt on stress is widely used in the analysis of fiber failures in fiber bundles (Newman and Phoenix, 2001).

As a first example, we considered a rod to which a constant uniaxial tensional stress $\sigma_0 > \sigma_y$ is applied instantaneously at $t = 0$ and held constant until the sample fails. The applicable kinetic equation for the rate of increase of damage with time is obtained from eqns [52] to [54] with the result

$$\frac{d\alpha}{dt}=\frac{1}{t_d}\left(\frac{\sigma_0}{\sigma_y}-1\right)^{\rho+2}\frac{1}{[1-\alpha(t)]^2} \quad [55]$$

Integrating with the initial condition $\alpha(0)=0$ gives

$$\alpha(t)=1-\left[1-3\frac{t}{t_d}\left(\frac{\sigma_0}{\sigma_y}-1\right)^{\rho+2}\right]^{1/3} \quad [56]$$

Substituting eqn [56] into eqn [52] gives the strain in the sample as a function of time t:

$$\frac{\epsilon(t)}{\epsilon_y}=1+\frac{(\sigma_0/\sigma_y-1)}{\left[1-3\frac{t}{t_d}(\sigma_0/\sigma_y-1)^{\rho+2}\right]^{1/3}} \quad [57]$$

Failure occurs at the time t_f when $\alpha\to 1$ ($\varepsilon\to\infty$). From eqn [57], this failure time is given by

$$t_f=\frac{t_d}{3}\left(\frac{\sigma_0}{\sigma_y}-1\right)^{-(\rho+2)} \quad [58]$$

The time to failure tends to infinity as a power law as the applied stress approaches the yield stress.

A particularly interesting set of experiments on brittle failure were carried out by Guarino et al. (1999). These authors studied the failure of circular panels (222 mm diameter, 3–5 mm thickness) of chipboard panels. A differential pressure was applied rapidly across a panel, and it was held constant until failure occurred. For these relatively thin panels, bending stresses were negligible and the panels failed under tension (a mode I fracture). Initially, the microcracks appeared to be randomly distributed across the panel; as the number of microcracks increased, they tended to localize and coalesce in the region where the final rupture occurred. The times to failure t_f of these chipboard panels were found to depend systematically on the applied differential pressure P. Taking a yield pressure (stress) $P_y=0.038$ MPa, their results correlate very well with eqn [57] taking $p=0.25$ and $t_d=168$ s. Shcherbakov et al. (2005a) and Ben-Zion and Lyakhovsky (2002) showed that the damage mechanic solution given in eqn [57] provides an explanation for the power-law increase in Benioff strain associated with AMR.

As our final example, we will consider a rod to which a constant uniaxial tensional strain $\varepsilon_0>\varepsilon_y$ has been applied instantaneously at $t=0$ and is held constant. The applicable equation for the rate of increase of damage with time is obtained from eqns [52] to [55] with the result

$$\frac{d\alpha}{dt}=\frac{1}{t_d}\left(\frac{\epsilon_0}{\epsilon_y}-1\right)^{\rho+2}[1-\alpha(t)]^{\rho} \quad [59]$$

Integrating with the initial condition $\alpha(0)=0$, we find

$$\alpha(t)=1-\left[1+(\rho-1)\frac{t}{t_d}\left(\frac{\epsilon_0}{\epsilon_y}-1\right)^{\rho+2}\right]^{-(1/\rho-1)} \quad [60]$$

This result is valid as long as $p>1$. The damage increases monotonically with time, and as $t\to\infty$, the maximum damage is $\alpha(\infty)=1$. Using eqn [60] with eqn [52], one gets the stress relaxation in the material as a function of time t

$$\frac{\sigma(t)}{\sigma_y}=1+\left(\frac{\epsilon_0}{\epsilon_y}-1\right)\left[1+(\rho-1)\frac{t}{t_d}\left(\frac{\epsilon_0}{\epsilon_y}-1\right)^{\rho+2}\right]^{-(1/\rho-1)} \quad [61]$$

At $t=0$, we have linear elasticity corresponding to $\alpha=0$. In the limit $t\to\infty$, the stress relaxes to the yield stress $\sigma(\infty)=\sigma_y$ below which no further damage can occur.

In order to relate our continuum damage mechanics model to aftershocks, we determine the rate of energy release in the relaxation process considered in the preceding text. Since the strain is constant during the stress relaxation, no work is done on the sample. We hypothesize that if the applied strain (stress) is instantaneously removed during the relaxation process, the sample will return to a state of zero stress and strain following a linear stress–strain path with slope $E_0(1-\alpha)$ to stress σ_y and a path with slope E_0 to zero stress. We assume that the difference between the energy added e_{YA} and the energy recovered e_{aY} is lost in aftershocks and find that this energy e_{as} is given by

$$e_{as}=e_{YA}-e_{aY}=\frac{1}{2}E_0(\epsilon_0-\epsilon_y)^2\alpha(t) \quad [62]$$

The rate of energy release is obtained by substituting eqn [59] into eqn [62] and taking the time derivative with the result

$$\frac{de_{as}}{dt}=\frac{E_0\epsilon_y^2}{2t_d}\left(\frac{\epsilon_0}{\epsilon_y}-1\right)^{\rho+4}\left[1+(\rho-1)\left(\frac{\epsilon_0}{\epsilon_y}-1\right)^{\rho+2}\left(\frac{t}{t_d}\right)\right]^{-(\rho/\rho-1)} \quad [63]$$

The total energy of aftershocks e_{ast} is obtained by substituting $\alpha=1$ into eqn [62] with the result

$$e_{ast}=\frac{1}{2}E_0(\epsilon_0-\epsilon_y)^2 \quad [64]$$

Combining eqns [63] and [64] gives

$$\frac{1}{e_{ast}}\frac{de_{as}}{dt}=\frac{\frac{1}{t_d}(\epsilon_0/\epsilon_y-1)^{\rho+2}}{\left[1+(\rho-1)(\epsilon_0/\epsilon_y-1)^{\rho+2}\left(\frac{t}{t_d}\right)\right]^{(\rho/\rho-1)}} \quad [65]$$

This result is clearly similar to the aftershock relation given in eqn [7]. To demonstrate this further, let us make the substitutions:

$$\rho=\frac{p}{p-1} \quad [66]$$

$$c=\frac{t_d}{(\rho-1)(\epsilon_0/\epsilon_y-1)^{\rho+2}} \quad [67]$$

Substitution of eqns [66] and [67] into eqn [65] gives

$$\frac{1}{e_{ast}}\frac{de_{as}}{dt}=\frac{p-1}{c}\frac{1}{(1+t/c)^p} \quad [68]$$

This damage mechanics result is identical in form to the modified Omori law for aftershocks given in eqn [7].

We next consider a specific example. In Figure 3, the rate of occurrence of aftershocks with $m\geq 2.0$ following the $m_{ms}=6.0$ Parkfield (California) earthquake, 28 September 2004, is given as a function of the time after the earthquake. Also shown is the correlation with eqn [7] taking $\tau=43.7$ s, $c=0.0199$ days, and $p=1.09$. From eqn [66], we find that the corresponding power-law exponent is $\rho=12.0$. Further assuming that $\varepsilon_0/\varepsilon_y=1.5$, we find from eqn [67] that the damage time is $t_d=0.01$ s.

4.24.4.5 Epidemic-Type Aftershock Sequence Models

It is clear that aftershocks play an important role in regional seismicity. One approach to assessing this role that has received considerable attention is the epidemic-type aftershock sequence (ETAS) branching model (Kagan and Knopoff, 1981; Ogata, 1988). This model recognizes that each primary aftershock will generate its own family of secondary aftershocks, the second-order aftershocks will generate third-order aftershocks, and so forth. This is a stochastic model that utilizes the GR frequency–magnitude distribution, the modified Omori law for temporal decay distribution, and a prescribed spatial distribution to create synthetic catalogs. Of particular importance is the prescribed branching ratio. The number of daughter earthquakes N_d generated by a parent earthquake with magnitude m_p is given by

$$N_d = g10^{\gamma(m_p - m_{min})} \quad [69]$$

where g and γ are specified constants and m_{min} is the magnitude of the smallest aftershock (daughter earthquake) that can be generated.

This approach was first given by Kagan and Knopoff (1981) and in a more comprehensive form by Ogata (1988). Further refinement has been given by Ogata (1992, 1999b, 2001, 2004), Ogata et al. (1993, 2003), Guo and Ogata (1997), Helmstetter and Sornette (2002a,b, 2003a,b), Helmstetter et al. (2003a,b), Sornette and Helmstetter (2002), Saichev and Sornette (2004, 2005), Saichev et al. (2005), and Sornette and Werner (2005a,b). In this model, each event is capable of producing secondary aftershock sequences and can be considered simultaneously as a foreshock, mainshock, or aftershock. The resulting aftershock sequence is a combined effect of many aftershock sequences produced by each earthquake.

The BASS model was introduced as the self-similar limit of the ETAS model (Holliday et al., 2008a,b; Lennartz et al., 2008, 2011; Turcotte et al., 2007). The primary difference between these models is that the BASS model incorporates the modified form of Båth's law given in eqn [4] instead of the branching ratio constant given in eqn [69]. The ETAS model reduces to the BASS model taking

$$\alpha = b \quad [70]$$

$$K = 10^{-b\Delta m^*} \quad [71]$$

where b is the b-value given in eqn [1] and Δm^* has been defined in eqn [4]. The BASS model limit eliminates the dependence on a specified minimum magnitude aftershock m_{min}.

The BASS model has several advantages over the generalized form of ETAS. The first is that Båth's law is automatically satisfied. The second is the elimination of a specified minimum aftershock magnitude m_{min}. In addition, BASS simulations generate aftershock statistics in agreement with simulations (Yoder et al., 2013). First, the frequency of foreshock occurrence scales as the exponential of the magnitude difference between the mainshock and the foreshock, inverse GR scaling. Second, the rate of foreshock occurrence decays inversely with the time difference between the foreshock and the mainshock, reverse Omori scaling.

4.24.5 Discussion

4.24.5.1 Overview

Earthquakes and seismicity are clearly complex phenomena. Yet, within this complexity, there are robust scaling laws. Both GR frequency–magnitude scaling and the Omori law for the temporal decay of aftershocks were introduced empirically. It can be shown that GR scaling is equivalent to fractal relationship between the number of earthquakes and their rupture area. This dependence is scale-invariant, it is robust, but do we understand it? One approach is to directly associate this scaling with the power-law slip-event scaling obtained in slider-block models. But in these models, an individual slider block can participate in events of all sizes. This does not seem to be the case for earthquakes on faults: big faults appear to have large earthquakes and small faults, small earthquakes. Thus, the GR scaling may be the consequence of a fractal distribution of fault sizes. This in turn can be attributed to scale-invariant comminution of the Earth's brittle crust in active tectonic regions. The established scaling laws and the statistical analysis of seismicity are the part of successful developments in statistical seismology (Vere-Jones, 2009, 2010).

The Omori law also appears to be universally applicable, but why? Why do aftershocks occur? A simple explanation is that some regions experience an increase in stress during a mainshock, although the average stress level decreases. Aftershocks relieve the excess stress in these regions. But why the time delay? A number of explanations have been given as discussed in **Sections 4.24.2–4.24.2.3**. Both damage mechanics and fiber-bundle models incorporate time delays for brittle failure on an empirical basis. But the fundamental physics of these time delays remains controversial. There are certainly similarities between these times and those associated with phase changes, for example, the explosive boiling of superheated water.

Another important question concerns the relative roles of aftershocks and 'characteristic earthquakes.' The ETAS model is an interesting approach to seismicity. GR and the Omori laws are assumed and earthquakes are generated randomly from these distributions. Because of this randomness, later earthquakes can be larger than the initial 'mainshock.' The original mainshock becomes a foreshock. But there is certainly a role for characteristic earthquakes and a tectonic drive. It is clearly desirable to obtain a better understanding of the relative roles of characteristic earthquakes and aftershocks.

4.24.5.2 Earthquake Forecasting (Prediction)

A fundamental question is whether forecasts of the time and location of future earthquakes can be accurately made (Tiampo and Shcherbakov, 2012). It is accepted that long-term hazard maps of the expected rate of occurrence of earthquakes are reasonably accurate. But is it possible to do better? Are there precursory phenomena that will allow earthquakes to be forecast?

It is actually quite surprising that short-term local precursory phenomena are not seen. Prior to a volcanic eruption, increases in regional seismicity and surface movements are generally observed (Chastin and Main, 2003). For a fault system, the stress gradually increases until it reaches the frictional strength of the fault and a rupture is initiated. It is certainly reasonable to hypothesize that the stress increase would cause increases in background seismicity and aseismic slip. In order to test this hypothesis, the Parkfield Earthquake Prediction Experiment was initiated in 1985. The expected Parkfield earthquake occurred beneath the heavily instrumented region on

28 September 2004. No local precursory changes were observed (Bakun et al., 2005).

In the absence of local precursory signals, the next question is whether broader anomalies develop and in particular whether there is anomalous seismic activity. One approach to earthquake forecasting is to use the rate of occurrence of small earthquakes in the past. This type of forecast is referred to as a 'relative intensity' (RI) forecast (Holliday et al., 2005, 2006; Shcherbakov et al., 2010). In such a forecast, the study region is tiled with boxes, that is, $0.1^\circ \times 0.1^\circ$. The number of earthquakes with magnitudes greater than a specified value in each box is determined over a specified time period. The RI score for each box is then computed as the total number of earthquakes in the box in the time period divided by the value for the box having the largest value. A threshold value in the interval [0–1] is then selected. Larger earthquakes are then forecast to occur only in boxes having an RI value larger than the threshold. The remaining boxes with RI scores smaller than the threshold represent sites at which larger earthquakes are forecast to not occur. The physical justification for this type of forecast is that large earthquakes are considered most likely to occur at sites of high seismic activity of small events in the past, that is, large 'a'-values in the GR scaling given in one.

A second approach to earthquake forecasting is based on the pattern informatics (PI) method, which quantifies temporal variations in seismicity (Holliday et al., 2005; Rundle et al., 2002b; Tiampo and Shcherbakov, 2012; Tiampo et al., 2002a,b,c). This approach is similar to the RI method. Again, the study region is tiled with boxes. Three time intervals are considered. The number of earthquakes per year in a box with magnitudes greater than a specified value is determined first for a reference interval t_0–t_2. Then the earthquake rates for a change interval t_1 to t_2 are determined. The changes in rates between the two intervals are quantified and the boxes with maximum changes are determined. These boxes ('hotspots') are the regions where earthquakes are forecast to occur during a forecast period t_2 to t_3. The PI method was first applied to seismicity in southern California and adjacent regions (32° to 37° N lat, 238° to 245° E long). This region was divided into a grid of 3500 boxes with $\Delta x = 0.1^\circ$ (11 km). Consistent with the sensitivity of the southern California seismic network, the lower magnitude cutoff was taken to be $m = 3.0$. The initial time was $t_0 = 1932$, the change interval was from $t_1 = 1990$ to $t_2 = 2000$, and the forecast interval was from $t_2 = 2000$ to $t_3 = 2010$. The initial studies for California were published in 2002 (Rundle et al., 2002b). This forecast was considered to be valid for the interval from 2000 to 2010 and would be applicable for earthquakes with $m = 5.0$ and larger. Since 1 January 2000, 18 earthquakes with $m \geq 5.0$ have occurred in the test region. We consider the forecast to be successful if the epicenter of the earthquake lies within a hotspot box or in one of the eight adjoining boxes. Sixteen of the eighteen earthquakes were successfully forecast.

A series of statistical algorithms to make intermediate-term earthquake predictions have been developed by a Russian group under the direction of V. I. Keilis-Borok using pattern recognition techniques (Keilis-Borok, 1990, 2002). Seismicity in circular regions with diameters of 500 km was analyzed. Based primarily on seismic activation, earthquake alarms were issued for one or more regions, with the alarms generally lasting for 3 years. Alarms have been issued regularly since the mid-1980s and scored two notable successes: the prediction of the 1988 Armenian earthquake and the 1989 Loma Prieta earthquake. While a reasonably high success rate has been achieved, there have been some notable misses including the recent $m = 9.3$ Sumatra and $m = 8.1$ Macquarie Island earthquakes.

More recently, this group has used chains of premonitory earthquakes as the basis for issuing alarms (Keilis-Borok et al., 2004; Shebalin et al., 2004). This method successfully predicted the $m = 6.5$, 22 December 2003 San Simeon (California) earthquake and the $m = 8.1$, 25 September 2003 Tokachi-oki (Japan) earthquake with lead times of 6 and 7 months, respectively. However, an alarm issued for southern California, valid during the spring and summer of 2004, was a false alarm. Both the PI method and the Russian forecast algorithms depend heavily on variations in seismicity prior to an earthquake. Thus, they may be related to AMR as discussed in **Sections 4.24.2–4.24.2.4**.

In order to test earthquake forecasts, it is clearly desirable to use prospective forecasts of future earthquakes. In order for a prospective test to be useful, it should be carried out in a reasonable length of time, say, 5 years, and a reasonable number of earthquakes should be expected to occur. In order to meet these criteria, the Regional Earthquake Likelihood Models test was developed and carried out by the Southern California Earthquake Center. The test regions were California and adjacent regions. Forecasts were solicited for $M \geq 5.0$ earthquakes during the period 2006–2010. The test region was divided into $0.1^\circ \times 0.1^\circ$ spatial cells. Participants were required to submit the number of earthquakes expected to occur in each cell. Six forecasts were submitted. Details of these forecasts have been given by Field (2007).

During the test period, there were 31 earthquakes that occurred in 22 test cells. The major earthquake that occurred was the $m = 7.2$ El Mayor–Cucapah earthquake on 4 April 2010. Six of the test earthquakes were classified as aftershocks of this earthquake, and seven of the test earthquakes constituted a precursory swarm, possibly a precursory activation. There are many alternative ways to evaluate the relative success of the submitted forecasts. Some of these have been discussed by Lee et al. (2011) and Sachs et al. (2012a,b). The three best forecasts had forecast probabilities of occurrence $\mathrm{Pr}_{\mathrm{forecast}} = 1.44 \times 10^{-3}$, 1.38×10^{-3}, and 1.13×10^{-3} for the 22 cells in which earthquakes occurred. A perfect forecast would have equal probabilities $\mathrm{Pr}_{\mathrm{forecast}} = 22^{-1} = 4.5 \times 10^{-2}$ for each cell in which an earthquake occurred and $\mathrm{Pr}_{\mathrm{forecast}} = 0.0$ in all other cells. A random forecast would give equal probabilities to the 7682 test cells, that is, $\mathrm{Pr}_{\mathrm{forecast}} = 7682^{-1} = 1.30 \times 10^{-4}$. Thus, the best forecasts were about a factor of 10 better than a random (no skill) forecast but were about a factor of 30 worse than a perfect forecast. These forecasts were based on extrapolations of the occurrence of past seismicity, that is, modifications of the RI method. A comprehensive review of seismicity-based earthquake forecasting techniques has been given by Tiampo and Shcherbakov (2012).

4.24.5.3 Future Directions

Many aspects of seismology exhibit complexity as has been discussed in previous sections. This is an active area of research in both the geophysical and statistical physics communities.

Although much progress has been made, many questions remain. Relevant areas include scaling laws, temporal and spatial correlations, critical phenomena, and nucleation.

Several scaling laws are now widely accepted. These include the GR frequency–magnitude scaling, the Omori law for the decay of aftershock activity, and Båth's law relating the magnitude of the largest aftershock to the magnitude of the mainshock. Other scaling laws are more controversial and require further studies. These include the power-law behavior of AMR, the Weibull distribution of recurrence times between characteristic earthquakes, and the nonhomogeneous Poisson distribution of interoccurrence times between aftershocks. The relative role of aftershocks and characteristic earthquakes needs to be clarified.

Models relevant to earthquakes and complexity are at an early stage of development. Slider-block models have certainly played a role but are clearly only weakly related to distributed seismicity. Damage mechanics also plays a role in understanding the complex behavior of brittle materials. Realistic simulations of distributed seismicity are just beginning to be developed. A major objective of these models is to provide estimates of the seismic hazard. A fundamental question is whether the future occurrence of earthquakes can be predicted. Certainly, the occurrence of small earthquakes is routinely utilized to forecast the risk of large earthquakes. This extrapolation is based on the applicability of GR frequency–magnitude statistics (Kossobokov et al., 2000). But can seismicity patterns be used to reliably forecast future earthquakes in a timely manner? There is accumulating evidence that this may be possible.

Acknowledgments

This work has been supported by National Science Foundation Grant ATM-0327571 (DT). It has been also supported by the NSERC Discovery Grant 355632-2008 and CFI Grant 25816 (RS).

References

Abaimov SG, Tiampo KF, Turcotte DL, and Rundle JB (2009) Recurrent frequency-size distribution of characteristic events. *Nonlinear Processes in Geophysics* 16(2): 333–350.

Abaimov SG, Turcotte DL, and Rundle JB (2007a) Recurrence-time and frequency-slip statistics of slip events on the creeping section of the San Andreas Fault in central California. *Geophysical Journal International* 170(3): 1289–1299.

Abaimov SG, Turcotte DL, Shcherbakov R, and Rundle JB (2007b) Recurrence and interoccurrence behavior of self-organized complex phenomena. *Nonlinear Processes in Geophysics* 14(4): 455–464.

Abaimov SG, Turcotte DL, Shcherbakov R, et al. (2008) Earthquakes: Recurrence and interoccurrence times. *Pure and Applied Geophysics* 165(3–4): 777–795.

Abe S and Suzuki N (2003) Law for the distance between successive earthquakes. *Journal of Geophysical Research* 108(B2): 2113.

Abe S and Suzuki N (2004a) Scale-free network of earthquakes. *Europhysics Letters* 65(4): 581–586.

Abe S and Suzuki N (2004b) Small-world structure of earthquake network. *Physica A* 337(1–2): 357–362.

Aki K (1981) A probabilistic synthesis of precursory phenomena. In: Simpson DW and Richards PG (eds.) *Earthquake Prediction: An International Review*. *Maurice Ewing Series*, vol. 4, pp. 566–574. Washington, DC: AGU.

Altmann EG, da Silva EC, and Caldas IL (2004) Recurrence time statistics for finite size intervals. *Chaos* 14(4): 975–981.

Altmann EG and Kantz H (2005) Recurrence time analysis, long-term correlations, and extreme events. *Physical Review E* 71(5), 056106.

Avrami M (1940) Kinetics of phase change. II: Transformation-time relations for random distribution of nuclei. *Journal of Chemical Physics* 8(2), 212–224.

Baiesi M (2006) Scaling and precursor motifs in earthquake networks. *Physica A* 360(2): 534–542.

Baiesi M and Paczuski M (2004) Scale-free networks of earthquakes and aftershocks. *Physical Review E* 69(6): 066106.

Baiesi M and Paczuski M (2005) Complex networks of earthquakes and aftershocks. *Nonlinear Processes in Geophysics* 12(1): 1–11.

Bak P, Christensen K, Danon L, and Scanlon T (2002) Unified scaling law for earthquakes. *Physical Review Letters* 88(17): 178501.

Bak P and Tang C (1989) Earthquakes as a self-organized critical phenomenon. *Journal of Geophysical Research* 94(B11): 15635–15637.

Bak P, Tang C, and Wiesenfeld K (1988) Self-organized criticality. *Physical Review A* 38(1): 364–374.

Bakun WH, Aagaard B, Dost B, et al. (2005) Implications for prediction and hazard assessment from the 2004 Parkfield earthquake. *Nature* 437(7061): 969–974.

Barriere B and Turcotte DL (1991) A scale-invariant cellular-automata model for distributed seismicity. *Geophysical Research Letters* 18(11): 2011–2014.

Barriere B and Turcotte DL (1994) Seismicity and self-organized criticality. *Physical Review E* 49(2): 1151–1160.

Båth M (1965) Lateral inhomogeneities of the upper mantle. *Tectonophysics* 2: 483–514.

Ben-Zion Y (2008) Collective behavior of earthquakes and faults: Continuum-discrete transitions, progressive evolutionary changes, and different dynamic regimes. *Reviews of Geophysics* 46(4): RG4006.

Ben-Zion Y and Lyakhovsky V (2002) Accelerated seismic release and related aspects of seismicity patterns on earthquake faults. *Pure and Applied Geophysics* 159(10): 2385–2412.

Ben-Zion Y and Lyakhovsky V (2006) Analysis of aftershocks in a lithospheric model with seismogenic zone governed by damage rheology. *Geophysical Journal International* 165(1): 197–210.

Bhattacharya P, Phan M, and Shcherbakov R (2011) Statistical analysis of the 2002 Mw 7.9 Denali earthquake. *Bulletin of Seismological Society of America* 101(6): 2662–2674.

Bowman DD and King GCP (2001) Accelerating seismicity and stress accumulation before large earthquakes. *Geophysical Research Letters* 28(21): 4039–4042.

Bowman DD, Ouillon G, Sammis CG, Sornette A, and Sornette D (1998) An observational test of the critical earthquake concept. *Journal of Geophysical Research* 103(B10): 24359–24372.

Bowman DD and Sammis CG (2004) Intermittent criticality and the Gutenberg–Richter distribution. *Pure and Applied Geophysics* 161(9–10): 1945–1956.

Brehm DJ and Braile LW (1998) Intermediate-term earthquake prediction using precursory events in the New Madrid seismic zone. *Bulletin of Seismological Society of America* 88(2): 564–580.

Brehm DJ and Braile LW (1999a) Intermediate-term earthquake prediction using the modified time-to-failure method in southern California. *Bulletin of Seismological Society of America* 89(1): 275–293.

Brehm DJ and Braile LW (1999b) Refinement of the modified time-to-failure method for intermediate-term earthquake prediction. *Journal of Seismology* 3(2): 121–138.

Brown SR, Scholz CH, and Rundle JB (1991) A simplified spring-block model of earthquakes. *Geophysical Research Letters* 18(2): 215–218.

Bufe CG, Nishenko SP, and Varnes DJ (1994) Seismicity trends and potential for large earthquakes in the Alaska-Aleutian region. *Pure and Applied Geophysics* 142(1): 83–99.

Bufe CG and Varnes DJ (1993) Predictive modeling of the seismic cycle of the greater San Francisco Bay region. *Journal of Geophysical Research* 98(B6): 9871–9883.

Bunde A, Eichner JF, Havlin S, and Kantelhardt JW (2003) The effect of long-term correlations on the return periods of rare events. *Physica A* 330(1–2): 1–7.

Bunde A, Eichner JF, Havlin S, and Kantelhardt JW (2004) Return intervals of rare events in records with long-term persistence. *Physica A* 342(1–2): 308–314.

Burridge R and Knopoff L (1967) Model and theoretical seismicity. *Bulletin of Seismological Society of America* 57: 341–371.

Burroughs SM and Tebbens SF (2002) The upper-truncated power law applied to earthquake cumulative frequency-magnitude distributions: Evidence for a time-independent scaling parameter. *Bulletin of Seismological Society of America* 92(8): 2983–2993.

Carbone V, Sorriso-Valvo L, Harabaglia P, and Guerra I (2005) Unified scaling law for waiting times between seismic events. *Europhysics Letters* 71(6): 1036–1042.

Carlson JM and Langer JS (1989) Properties of earthquakes generated by fault dynamics. *Physical Review Letters* 62(22): 2632–2635.

Chastin SFM and Main IG (2003) Statistical analysis of daily seismic event rate as a precursor to volcanic eruptions. *Geophysical Research Letters* 30(13): 1671.

Coleman BD (1958) Statistics and time dependence of mechanical breakdown in fibers. *Journal of Applied Physics* 29: 968–983.

Console R, Lombardi AM, Murru M, and Rhoades D (2003a) Båth's law and the self-similarity of earthquakes. *Journal of Geophysical Research* 108(B2): 2128.

Console R and Murru M (2001) A simple and testable model for earthquake clustering. *Journal of Geophysical Research* 106(B5): 8699–8711.

Console R, Murru M, and Lombardi AM (2003b) Refining earthquake clustering models. *Journal of Geophysical Research* 108(B10): 2468.

Corral Á (2003) Local distributions and rate fluctuations in a unified scaling law for earthquakes. *Physical Review E* 68(3): 035102.

Corral Á (2004a) Long-term clustering, scaling, and universality in the temporal occurrence of earthquakes. *Physical Review Letters* 92(10): 108501.

Corral Á (2004b) Universal local versus unified global scaling laws in the statistics of seismicity. *Physica A* 340(4): 590–597.

Corral Á (2005a) Mixing of rescaled data and Bayesian inference for earthquake recurrence times. *Nonlinear Processes in Geophysics* 12(1): 89–100.

Corral Á (2005b) Time-decreasing hazard and increasing time until the next earthquake. *Physical Review E* 71(1): 017101.

Corral Á (2006a) Dependence of earthquake recurrence times and independence of magnitudes on seismicity history. *Tectonophysics* 424(3–4): 177–193.

Corral Á (2006b) Statistical features of earthquake temporal occurrence. In: Bhattacharyya P and Chakrabarti B (eds.) *Modelling Critical and Catastrophic Phenomena in Geoscience. Lecture Notes in Physics*, vol. 705, pp. 191–221. Berlin/Heidelberg: Springer.

Corral Á (2007) Structure of earthquake occurrence in space, time and magnitude. *Terra Nova* 19(5): 337–343.

Dahmen KA and Ben-Zion Y (2009) Physics of jerky motion in slowly driven magnetic and earthquake fault systems. In: Meyers RA (ed.) *Encyclopedia of Complexity and Systems Science*, pp. 5021–5037. New York: Springer.

Das S and Scholz CH (1981) Off-fault aftershock clusters caused by shear-stress increase. *Bulletin of Seismological Society of America* 71(5): 1669–1675.

Davidsen J and Goltz C (2004) Are seismic waiting time distributions universal? *Geophysical Research Letters* 31(21): L21612.

Davidsen J, Grassberger P, and Paczuski M (2006) Earthquake recurrence as a record breaking process. *Geophysical Research Letters* 33(11): L11304.

Davidsen J, Grassberger P, and Paczuski M (2008) Networks of recurrent events, a theory of records, and an application to finding causal signatures in seismicity. *Physical Review E* 77(6): 066104.

Davidsen J and Paczuski M (2005) Analysis of the spatial distribution between successive earthquakes. *Physical Review Letters* 94(4), 048501.

Davis PM, Jackson DD, and Kagan YY (1989) The longer it has been since the last earthquake, the longer the expected time till the next. *Bulletin of Seismological Society of America* 79(5): 1439–1456.

Dhar D (1990) Self-organized critical state of sandpile automaton models. *Physical Review Letters* 64(14): 1613–1616.

Dhar D (1999) The Abelian sandpile and related models. *Physica A* 263(1–4): 4–25.

Dieterich J (1994) A constitutive law for rate of earthquake production and its application to earthquake clustering. *Journal of Geophysical Research* 99(B2): 2601–2618.

Dieterich JH and Richards-Dinger KB (2010) Earthquake recurrence in simulated fault systems. *Pure and Applied Geophysics* 167(8–9): 1087–1104.

Dobrovolsky IP, Zubkov SI, and Miachkin VI (1979) Estimation of the size of earthquake preparation zones. *Pure and Applied Geophysics* 117(5): 1025–1044.

Drossel B and Schwabl F (1992) Self-organized critical forest-fire model. *Physical Review Letters* 69(11): 1629–1632.

Ellsworth WL, Matthews MV, Nadeau RM, Nishenko SP, Reasenberg PA, and Simpson RW (1999) A physically-based earthquake recurrence model for estimation of long-term earthquake probabilities. *Technical Report, Open-File Report 99-522*. USGS.

Enescu B, Mori J, Miyazawa M, and Kano Y (2009) Omori–Utsu law *c*-values associated with recent moderate earthquakes in Japan. *Bulletin of Seismological Society of America* 99(2A): 884–891.

Felzer KR, Abercrombie RE, and Ekstrom G (2003) Secondary aftershocks and their importance for aftershock forecasting. *Bulletin of Seismological Society of America* 93(4): 1433–1448.

Felzer KR, Abercrombie RE, and Ekstrom G (2004) A common origin for aftershocks, foreshocks, and multiplets. *Bulletin of Seismological Society of America* 94(1): 88–98.

Felzer KR, Becker TW, Abercrombie RE, Ekstrom G, and Rice JR (2002) Triggering of the 1999 M_w 7.1 Hector Mine earthquake by aftershocks of the 1992 M_w 7.3 Landers earthquake. *Journal of Geophysical Research* 107(B9): 2190.

Felzer KR and Brodsky EE (2006) Decay of aftershock density with distance indicates triggering by dynamic stress. *Nature* 441(7094): 735–738.

Field EH (2007) Overview of the working group for the development of regional earthquake likelihood models (RELM). *Seismological Research Letters* 78(1): 7–16.

Frohlich C and Davis SD (1990) Single-link cluster-analysis as a method to evaluate spatial and temporal properties of earthquake catalogs. *Geophysical Journal International* 100(1): 19–32.

Frohlich C and Davis SD (1993) Teleseismic *b*-values – Or, much ado about 1.0. *Journal of Geophysical Research* 98(B1): 631–644.

Gabrielov A, Newman WI, and Turcotte DL (1999) Exactly soluble hierarchical clustering model: Inverse cascades, self-similarity, and scaling. *Physical Review E* 60(5): 5293–5300.

Geller RJ (1997) Earthquake prediction: A critical review. *Geophysical Journal International* 131(3): 425–450.

Geller RJ, Jackson DD, Kagan YY, and Mulargia F (1997) Earthquakes cannot be predicted. *Science* 275(5306): 1616–1617.

Goes SDB and Ward SN (1994) Synthetic seismicity for the San Andreas Fault. *Annals of Geophysics* 37: 1495–1513.

Goltz C and Bose M (2002) Configurational entropy of critical earthquake populations. *Geophysical Research Letters* 29(20), 1990.

Goltz C, Turcotte DL, Abaimov SG, Nadeau RM, Uchida N, and Matsuzawa T (2009) Rescaled earthquake recurrence time statistics: Application to microrepeaters. *Geophysical Journal International* 176(1): 256–264.

Gross S and Rundle J (1998) A systematic test of time-to-failure analysis. *Geophysical Journal International* 133(1): 57–64.

Guarino A, Ciliberto S, and Garcimartin A (1999) Failure time and microcrack nucleation. *Europhysics Letters* 47(4): 456–461.

Guo ZQ and Ogata Y (1997) Statistical relations between the parameters of aftershocks in time, space, and magnitude. *Journal of Geophysical Research* 102(B2): 2857–2873.

Gutenberg B and Richter CF (1954) *Seismicity of the Earth and Associated Phenomenon*, 2nd edn. Princeton, NJ: Princeton University Press.

Hardebeck JL, Felzer KR, and Michael AJ (2008) Improved tests reveal that the accelerating moment release hypothesis is statistically insignificant. *Journal of Geophysical Research* 113(B8): B08310.

Helmstetter A, Kagan YY, and Jackson DD (2005) Importance of small earthquakes for stress transfers and earthquake triggering. *Journal of Geophysical Research* 110(B5): B05S08.

Helmstetter AS, Ouillon G, and Sornette D (2003a) Are aftershocks of large Californian earthquakes diffusing? *Journal of Geophysical Research* 108(B10): 2483.

Helmstetter A and Sornette D (2002a) Diffusion of epicenters of earthquake aftershocks, Omori's law, and generalized continuous-time random walk models. *Physical Review E* 66(6): 061104.

Helmstetter A and Sornette D (2002b) Subcritical and supercritical regimes in epidemic models of earthquake aftershocks. *Journal of Geophysical Research* 107(B10): 2237.

Helmstetter A and Sornette D (2003a) Båth's law derived from the Gutenberg–Richter law and from aftershock properties. *Geophysical Research Letters* 30(20): 2069.

Helmstetter A and Sornette D (2003b) Predictability in the epidemic-type aftershock sequence model of interacting triggered seismicity. *Journal of Geophysical Research* 108(B10): 2482.

Helmstetter A, Sornette D, and Grasso JR (2003b) Mainshocks are aftershocks of conditional foreshocks: How do foreshock statistical properties emerge from aftershock laws. *Journal of Geophysical Research* 108(B1): 2046.

Henderson JR, Main IG, Maclean C, and Norman MG (1994) A fracture-mechanical cellular-automaton model of seismicity. *Pure and Applied Geophysics* 142(3–4): 545–565.

Henley CL (1993) Statics of a "self-organized" percolation model. *Physical Review Letters* 71(17): 2741.

Hergarten S and Krenn R (2011) Synchronization and desynchronization in the Olami–Feder–Christensen earthquake model and potential implications for real seismicity. *Nonlinear Processes in Geophysics* 18(5): 635–642.

Hill DP, Reasenberg PA, Michael A, et al. (1993) Seismicity remotely triggered by the magnitude 7.3 Landers, California, earthquake. *Science* 260(5114): 1617–1623.

Holliday JR, Nanjo KZ, Tiampo KF, Rundle JB, and Turcotte DL (2005) Earthquake forecasting and its verification. *Nonlinear Processes in Geophysics* 12: 965–977.

Holliday JR, Rundle JB, Tiampo KF, and Turcotte DL (2006) Using earthquake intensities to forecast earthquake occurrence times. *Nonlinear Processes in Geophysics* 13(5): 585–593.

Holliday JR, Turcotte DL, and Rundle JB (2008a) A review of earthquake statistics: Fault and seismicity-based models. ETAS and BASS. *Pure and Applied Geophysics* 165(6): 1003–1024.

Holliday JR, Turcotte DL, and Rundle JB (2008b) Self-similar branching of aftershock sequences. *Physica A* 387(4): 933–943.
Huang Y, Saleur H, Sammis C, and Sornette D (1998) Precursors, aftershocks, criticality and self-organized criticality. *Europhysics Letters* 41(1): 43–48.
Huang J and Turcotte DL (1990) Are earthquakes an example of deterministic chaos. *Geophysical Research Letters* 17(3): 223–226.
Huang J and Turcotte DL (1992) Chaotic seismic faulting with a mass-spring model and velocity-weakening friction. *Pure and Applied Geophysics* 138(4): 569–589.
Huc M and Main IG (2003) Anomalous stress diffusion in earthquake triggering: Correlation length, time dependence, and directionality. *Journal of Geophysical Research* 108(B7): 2324.
Ivashkevich EV and Priezzhev VB (1998) Introduction to the sandpile model. *Physica A* 254(1–2): 97–116.
Jaumé SC (2000) Changes in earthquake size-frequency distributions underlying accelerating moment/energy release. In: *Geocomplexity and the Physics of Earthquakes*. *Geophysical Monograph Series*, vol. 120, pp. 199–210. Washington, DC: AGU.
Jaumé SC and Sykes LR (1999) Evolving towards a critical point: A review of accelerating seismic moment/energy release prior to large and great earthquakes. *Pure and Applied Geophysics* 155(2–4): 279–305.
Jensen HJ (1998) *Self-Organized Criticality: Emergent Complex Behavior in Physical and Biological Systems*. Cambridge: Cambridge University Press.
Jiang CS and Wu ZL (2005) Test of the preshock accelerating moment release (AMR) in the case of the 26 December 2004 Mw 9.0 Indonesia earthquake. *Bulletin of Seismological Society of America* 95(5): 2016–2025.
Jiang CS and Wu ZL (2006) Benioff strain release before earthquakes in China: Accelerating or not? *Pure and Applied Geophysics* 163(9): 1965–1976.
Kachanov LM (1986) *Introduction to Continuum Damage Mechanics*. Dordrecht: Martinus Nijhoff.
Kadanoff LP, Nagel SR, Wu L, and Zhou SM (1989) Scaling and universality in avalanches. *Physical Review A* 39(12): 6524–6537.
Kagan YY (2004) Short-term properties of earthquake catalogs and models of earthquake source. *Bulletin of Seismological Society of America* 94(4): 1207–1228.
Kagan YY and Knopoff L (1981) Stochastic synthesis of earthquake catalogs. *Journal of Geophysical Research* 86(NB4): 2853–2862.
Kanamori H and Anderson DL (1975) Theoretical basis of some empirical relations in seismology. *Bulletin of Seismological Society of America* 65(5): 1073–1095.
Kanamori H and Brodsky EE (2004) The physics of earthquakes. *Reports on Progress in Physics* 67(8): 1429–1496.
Kawamura H, Hatano T, Kato N, Biswas S, and Chakrabarti BK (2012) Statistical physics of fracture, friction, and earthquakes. *Reviews of Modern Physics* 84(2): 839–884.
Kawamura H, Yamamoto T, Kotani T, and Yoshino H (2010) Asperity characteristics of the Olami–Feder–Christensen model of earthquakes. *Physical Review E* 81(3): 031119.
Keilis-Borok VI (1990) The lithosphere of the earth as a nonlinear-system with implications for earthquake prediction. *Reviews of Geophysics* 28(1): 19–34.
Keilis-Borok V (2002) Earthquake prediction: State-of-the-art and emerging possibilities. *Annual Review of Earth and Planetary Sciences* 30: 1–33.
Keilis-Borok VI and Kossobokov VG (1990) Premonitory activation of earthquake flow – Algorithm M8. *Physics of the Earth and Planetary Interiors* 61(1–2): 73–83.
Keilis-Borok V, Shebalin P, Gabrielov A, and Turcotte D (2004) Reverse tracing of short-term earthquake precursors. *Physics of the Earth and Planetary Interiors* 145(1–4): 75–85.
King GCP and Bowman DD (2003) The evolution of regional seismicity between large earthquakes. *Journal of Geophysical Research* 108(B2): 2096.
Kisslinger C (1996) Aftershocks and fault-zone properties. *Advanced Geophysics*, vol. 38, pp. 1–36. San Diego, CA: Academic Press.
Kisslinger C and Jones LM (1991) Properties of aftershock sequences in Southern California. *Journal of Geophysical Research* 96(B7): 11947–11958.
Knopoff L (1996) Earthquake prediction: The scientific challenge. *Proceedings of the National Academy of Sciences of the United States of America* 93(9): 3719–3720.
Kossobokov VG, Keilis-Borok VI, Turcotte DL, and Malamud BD (2000) Implications of a statistical physics approach for earthquake hazard assessment and forecasting. *Pure and Applied Geophysics* 157(11–12): 2323–2349.
Kotani T, Yoshino H, and Kawamura H (2008) Periodicity and criticality in the Olami–Feder–Christensen model of earthquakes. *Physical Review E* 77(1): 010102.
Krajcinovic D (1996) *Damage Mechanics*. Amsterdam: Elsevier.
Kwiatek G, Plenkers K, Nakatani M, Yabe Y, Dresen G, and JAGUARS-Group (2010) Frequency-magnitude characteristics down to magnitude −4.4 for induced seismicity recorded at Mponeng gold mine, South Africa. *Bulletin of Seismological Society of America* 100(3): 1165–1173.
Lee YT, Turcotte DL, Holliday JR, et al. (2011) Results of the Regional Earthquake Likelihood Models (RELM) test of earthquake forecasts in California. *Proceedings of the National Academy of Sciences of the United States of America* 108(40): 16533–16538.
Lennartz S, Bunde A, and Turcotte DL (2008) Missing data in aftershock sequences: Explaining the deviations from scaling laws. *Physical Review E* 78(4): 041115.
Lennartz S, Bunde A, and Turcotte DL (2011) Modelling seismic catalogues by cascade models: Do we need long-term magnitude correlations? *Geophysical Journal International* 184(3): 1214–1222.
Lindman M, Jonsdottir K, Roberts R, Lund B, and Bodvarsson R (2005) Earthquakes descaled: On waiting time distributions and scaling laws. *Physical Review Letters* 94(10): 108501.
Livina VN, Havlin S, and Bunde A (2005a) Memory in the occurrence of earthquakes. *Physical Review Letters* 95(20): 208501.
Livina V, Tuzov S, Havlin S, and Bunde A (2005b) Recurrence intervals between earthquakes strongly depend on history. *Physica A* 348: 591–595.
Lorenz EN (1963) Deterministic nonperiodic flow. *Journal of the Atmospheric Sciences* 20: 130–141.
Lyakhovsky V and Ben-Zion Y (2008) Scaling relations of earthquakes and aseismic deformation in a damage rheology model. *Geophysical Journal International* 172(2): 651–662.
Lyakhovsky V, Ben-Zion Y, and Agnon A (1997) Distributed damage, faulting, and friction. *Journal of Geophysical Research* 102(B12): 27635–27649.
Lyakhovsky V, Ben-Zion Y, and Agnon A (2001) Earthquake cycle, fault zones, and seismicity patterns in a rheologically layered lithosphere. *Journal of Geophysical Research* 106(B3): 4103–4120.
Ma S-K (1976) *Modern Theory of Critical Phenomena*. New York: W. A. Benjamin Reading.
Main I (1996) Statistical physics, seismogenesis, and seismic hazard. *Reviews of Geophysics* 34(4): 433–462.
Main IG (1999) Applicability of time-to-failure analysis to accelerated strain before earthquakes and volcanic eruptions. *Geophysical Journal International* 139(3): F1–F6.
Main IG (2000) A damage mechanics model for power-law creep and earthquake aftershock and foreshock sequences. *Geophysical Journal International* 142(1): 151–161.
Main IG and Naylor M (2008) Maximum entropy production and earthquake dynamics. *Geophysical Research Letters* 35(19), L19311.
Main IG and Naylor M (2010) Entropy production and self-organized (sub)criticality in earthquake dynamics. *Philosophical Transactions of the Royal Society A* 368(1910): 131–144.
Mandelbrot BB (1967) How long is the coast of Britain? Statistical self-similarity and fractional dimension. *Science* 156(3775): 636–638.
Matthews MV, Ellsworth WL, and Reasenberg PA (2002) A Brownian model for recurrent earthquakes. *Bulletin of Seismological Society of America* 92(6): 2233–2250.
McKernon C and Main IG (2005) Regional variations in the diffusion of triggered seismicity. *Journal of Geophysical Research* 110(B5): B05S05.
Meeker WA and Escobar LA (1998) *Statistical Methods for Reliability Data*. New York: Wiley.
Mignan A (2008) The stress accumulation model: Accelerating moment release and seismic hazard. *Advances in Geophysics* 49: 67–201.
Mignan A (2011) Retrospective on the accelerating seismic release (ASR) hypothesis: Controversy and new horizons. *Tectonophysics* 505(1–4): 1–16.
Mignan A, Bowman DD, and King GCP (2006) An observational test of the origin of accelerating moment release before large earthquakes. *Journal of Geophysical Research* 111: B11304.
Mignan A, King GCP, and Bowman D (2007) A mathematical formulation of accelerating moment release based on the stress accumulation model. *Journal of Geophysical Research* 112: B07308.
Nakanishi H (1990) Cellular-automaton model of earthquakes with deterministic dynamics. *Physical Review A* 41(12): 7086–7089.
Nanjo KZ, Enescu B, Shcherbakov R, Turcotte DL, Iwata T, and Ogata Y (2007) Decay of aftershock activity for Japanese earthquakes. *Journal of Geophysical Research* 112(B8): B08309.
Narteau C, Byrdina S, Shebalin P, and Schorlemmer D (2009) Common dependence on stress for the two fundamental laws of statistical seismology. *Nature* 462(7273): 642–645.
Naylor M and Main IG (2008) Cell scale self-organisation in the OFC model for earthquake dynamics. *European Physical Journal B* 64(1): 139–146.
Newman WI and Phoenix SL (2001) Time-dependent fiber bundles with local load sharing. *Physical Review E* 63(2): 021507.
Newman WI and Turcotte DL (2002) A simple model for the earthquake cycle combining self-organized complexity with critical point behavior. *Nonlinear Processes in Geophysics* 9(5–6): 453–461.

Nicolis G and Nicolis C (2012) *Foundations of Complex Systems: Emergence, Information and Prediction*, 2nd edn. Singapore: World Scientific Publishing.

Ogata Y (1988) Statistical-models for earthquake occurrences and residual analysis for point-processes. *Journal of the American Statistical Association* 83(401): 9–27.

Ogata Y (1992) Detection of precursory relative quiescence before great earthquakes through a statistical-model. *Journal of Geophysical Research* 97(B13): 19845–19871.

Ogata Y (1999a) Estimating the hazard of rupture using uncertain occurrence times of paleoearthquakes. *Journal of Geophysical Research* 104(B8): 17995–18014.

Ogata Y (1999b) Seismicity analysis through point-process modeling: A review. *Pure and Applied Geophysics* 155(2–4): 471–507.

Ogata Y (2001) Increased probability of large earthquakes near aftershock regions with relative quiescence. *Journal of Geophysical Research* 106(B5): 8729–8744.

Ogata Y (2004) Space-time model for regional seismicity and detection of crustal stress changes. *Journal of Geophysical Research* 109(B6): B06308.

Ogata Y, Jones LM, and Toda S (2003) When and where the aftershock activity was depressed: Contrasting decay patterns of the proximate large earthquakes in southern California. *Journal of Geophysical Research* 108(B6): 2318.

Ogata Y, Matsuura RS, and Katsura K (1993) Fast likelihood computation of epidemic type aftershock-sequence model. *Geophysical Research Letters* 20(19): 2143–2146.

Ojala IO, Ngwenya BT, Main IG, and Elphick SC (2003) Correlation of microseismic and chemical properties of brittle deformation in locharbriggs sandstone. *Journal of Geophysical Research* 108(B5): 2268.

Olami Z, Feder HJS, and Christensen K (1992) Self-organized criticality in a continuous, nonconservative cellular automaton modeling earthquakes. *Physical Review Letters* 68(8): 1244–1247.

Peng ZG, Vidale JE, and Houston H (2006) Anomalous early aftershock decay rate of the 2004 Mw 6.0 Parkfield, California, earthquake. *Geophysical Research Letters* 33(17): L17307.

Peng ZG, Vidale JE, Ishii M, and Helmstetter A (2007) Seismicity rate immediately before and after main shock rupture from high-frequency waveforms in Japan. *Journal of Geophysical Research* 112(B3), B03306.

Pollitz FF (2011) Epistemic uncertainty in California-wide synthetic seismicity simulations. *Bulletin of Seismological Society of America* 101(5): 2481–2498.

Priezzhev VB, Ktitarev DV, and Ivashkevich EV (1996) Formation of avalanches and critical exponents in an Abelian sandpile model. *Physical Review Letters* 76(12): 2093–2096.

Reasenberg PA and Jones LM (1989) Earthquake hazard after a mainshock in California. *Science* 243(4895): 1173–1176.

Robertson MC, Sammis CG, Sahimi M, and Martin AJ (1995) Fractal analysis of 3-dimensional spatial distributions of earthquakes with a percolation interpretation. *Journal of Geophysical Research* 100(B1): 609–620.

Robinson R (2000) A test of the precursory accelerating moment release model on some recent New Zealand earthquakes. *Geophysical Journal International* 140(3): 568–576.

Robinson R and Benites R (1995) Synthetic seismicity models of multiple interacting faults. *Journal of Geophysical Research* 100(B9): 18229–18238.

Robinson R and Benites R (1996) Synthetic seismicity models for the Wellington region, New Zealand: Implications for the temporal distribution of large events. *Journal of Geophysical Research* 101(B12): 27833–27844.

Robinson R, Zhou SY, Johnston S, and Vere-Jones D (2005) Precursory accelerating seismic moment release (AMR) in a synthetic seismicity catalog: A preliminary study. *Geophysical Research Letters* 32(7): L07309.

Rundle JB (1988) A physical model for earthquakes: 1. Fluctuations and interactions. *Journal of Geophysical Research* 93(B6): 6237–6254.

Rundle JB (1989a) Derivation of the complete Gutenberg–Richter magnitude–frequency relation using the principle of scale invariance. *Journal of Geophysical Research* 94(B9): 12337–12342.

Rundle JB (1989b) A physical model for earthquakes: 3. Thermodynamical approach and its relation to nonclassical theories of nucleation. *Journal of Geophysical Research* 94(B3): 2839–2855.

Rundle JB and Klein W (1993) Scaling and critical phenomena in a cellular automation slider-block model for earthquakes. *Journal of Statistical Physics* 72(1–2): 405–412.

Rundle JB, Klein W, and Gross S (1999) Physical basis for statistical patterns in complex earthquake populations: Models, predictions and tests. *Pure and Applied Geophysics* 155(2–4): 575–607.

Rundle J, Klein W, Turcotte DL, and Malamud BD (2000) Precursory seismic activation and critical-point phenomena. *Pure and Applied Geophysics* 157(11–12): 2165–2182.

Rundle JB, Rundle PB, Donnellan A, and Fox G (2004) Gutenberg–Richter statistics in topologically realistic system-level earthquake stress-evolution simulations. *Earth, Planets and Space* 56(8): 761–771.

Rundle JB, Rundle PB, Donnellan A, et al. (2005) A simulation-based approach to forecasting the next great San Francisco earthquake. *Proceedings of the National Academy of Sciences of the United States of America* 102(43): 15363–15367.

Rundle JB, Rundle PB, Klein W, et al. (2002a) GEM plate boundary simulations for the plate boundary observatory: A program for understanding the physics of earthquakes on complex fault networks via observations, theory and numerical simulation. *Pure and Applied Geophysics* 159(10): 2357–2381.

Rundle PB, Rundle JB, Tiampo KF, Donnellan A, and Turcotte DL (2006) Virtual California: Fault model, frictional parameters, applications. *Pure and Applied Geophysics* 163(9): 1819–1846.

Rundle PB, Rundle JB, Tiampo KF, Martins JSS, McGinnis S, and Klein W (2001) Nonlinear network dynamics on earthquake fault systems. *Physical Review Letters* 8714(14): 148501.

Rundle JB, Tiampo KF, Klein W, and Martins JSS (2002b) Self-organization in leaky threshold systems: The influence of near-mean field dynamics and its implications for earthquakes, neurobiology, and forecasting. *Proceedings of the National Academy of Sciences of the United States of America* 99(supplement 1): 2514–2521.

Rundle JB, Turcotte DL, Shcherbakov R, Klein W, and Sammis C (2003) Statistical physics approach to understanding the multiscale dynamics of earthquake fault systems. *Reviews of Geophysics* 41(4), 1019.

Sachs MK, Lee YT, Turcotte DL, Holliday JR, and Rundle JB (2012a) Evaluating RELM test results. *International Journal of Geophysics* 543482.

Sachs MK, Lee YT, Turcotte DL, Holliday JR, and Rundle JB (2012b) Implications of the Regional Earthquake Likelihood Models test of earthquake forecasts in California. *Research in Geophysics* 2: e10.

Saichev A, Helmstetter A, and Sornette D (2005) Power-law distributions of offspring and generation numbers in branching models of earthquake triggering. *Pure and Applied Geophysics* 162(6–7): 1113–1134.

Saichev A and Sornette D (2004) Anomalous power law distribution of total lifetimes of branching processes: Application to earthquake aftershock sequences. *Physical Review E* 70(4): 046123.

Saichev A and Sornette D (2005) Distribution of the largest aftershocks in branching models of triggered seismicity: Theory of the universal Båth law. *Physical Review E* 71(5): 056127.

Saleur H, Sammis CG, and Sornette D (1996a) Discrete scale invariance, complex fractal dimensions, and log-periodic fluctuations in seismicity. *Journal of Geophysical Research* 101(B8): 17661–17677.

Saleur H, Sammis CG, and Sornette D (1996b) Renormalization group theory of earthquakes. *Nonlinear Processes in Geophysics* 3: 102–109.

Sammis CG, Bowman DD, and King G (2004) Anomalous seismicity and accelerating moment release preceding the 2001 and 2002 earthquakes in northern Baja California, Mexico. *Pure and Applied Geophysics* 161(11–12): 2369–2378.

Sammis CG and Sornette D (2002) Positive feedback, memory, and the predictability of earthquakes. *Proceedings of the National Academy of Sciences of the United States of America* 99(supplement 1): 2501–2508.

Sanchez R, Newman DE, and Carreras BA (2002) Waiting-time statistics of self-organized-criticality systems. *Physical Review Letters* 88(6): 068302.

Scholz CH (1991) Earthquakes and faulting: Self-organized critical phenomena with a characteristic dimension. In: Riste T and Sherrington D (eds.) *Spontaneous Formation of Space Time Structure and Criticality*, pp. 41–56. Norwell, MA: Kluwer Academic.

Scholz CH (2002) *The Mechanics of Earthquakes and Faulting*, 2nd edn. Cambridge: Cambridge University Press.

Shaw BE (1993) Generalized Omori law for aftershocks and foreshocks from a simple dynamics. *Geophysical Research Letters* 20(10): 907–910.

Shcherbakov R, Davidsen J, and Tiampo KF (2013a) Record-breaking avalanches in driven threshold systems. *Physical Review E* 87: 052811.

Shcherbakov R, Goda K, Ivanian A, and Atkinson GM (2013b) Aftershock statistics of major subduction earthquakes. *Bulletin of Seismological Society of America* 103: 3222–3234.

Shcherbakov R, Nguyen M, and Quigley M (2012) Statistical analysis of the 2010 Mw 7.1 Darfield earthquake aftershock sequence. *New Zealand Journal of Geology and Geophysics* 55(3): 305–311.

Shcherbakov R and Turcotte DL (2000) Self-organizing height-arrow model: Numerical and analytical results. *Physica A* 277(3–4): 274–292.

Shcherbakov R and Turcotte DL (2003) Damage and self-similarity in fracture. *Theoretical and Applied Fracture Mechanics* 39(3): 245–258.

Shcherbakov R and Turcotte DL (2004) A modified form of Båth's law. *Bulletin of Seismological Society of America* 94(5): 1968–1975.

Shcherbakov R, Turcotte DL, and Rundle JB (2004) A generalized Omori's law for earthquake aftershock decay. *Geophysical Research Letters* 31(11): L11613.

Shcherbakov R, Turcotte DL, and Rundle JB (2005a) Aftershock statistics. *Pure and Applied Geophysics* 162(6–7): 1051–1076.
Shcherbakov R, Turcotte DL, and Rundle JB (2006a) Scaling properties of the Parkfield aftershock sequence. *Bulletin of Seismological Society of America* 96(4B): S376–S384.
Shcherbakov R, Turcotte DL, Rundle JB, Tiampo KF, and Holliday JR (2010) Forecasting the locations of future large earthquakes: An analysis and verification. *Pure and Applied Geophysics* 167(6–7): 743–749.
Shcherbakov R, Van Aalsburg J, Rundle JB, and Turcotte DL (2006b) Correlations in aftershock and seismicity patterns. *Tectonophysics* 413(1–2): 53–62.
Shcherbakov R, Yakovlev G, Turcotte DL, and Rundle JB (2005b) Model for the distribution of aftershock interoccurrence times. *Physical Review Letters* 95: 218501.
Shebalin P, Keilis-Borok V, Zaliapin I, Uyeda S, Nagao T, and Tsybin N (2004) Advance short-term prediction of the large Tokachi-oki earthquake, September 25, 2003, M = 8.1: A case history. *Earth, Planets and Space* 56(8): 715–724.
Sornette D (2006) *Critical Phenomena in Natural Sciences: Chaos, Fractals, Selforganization, and Disorder: Concepts and Tools*. Springer Series in Synergetics, 2nd edn. Berlin/New York: Springer.
Sornette D and Helmstetter A (2002) Occurrence of finite-time singularities in epidemic models of rupture, earthquakes, and starquakes. *Physical Review Letters* 89(15): 158501.
Sornette D and Knopoff L (1997) The paradox of the expected time until the next earthquake. *Bulletin of Seismological Society of America* 87(4): 789–798.
Sornette D, Knopoff L, Kagan YY, and Vanneste C (1996) Rank-ordering statistics of extreme events: Application to the distribution of large earthquakes. *Journal of Geophysical Research* 101(B6): 13883–13893.
Sornette D and Sammis CG (1995) Complex critical exponents from renormalization-group theory of earthquakes – Implications for earthquake predictions. *Journal de Physique I* 5(5): 607–619.
Sornette D and Werner MJ (2005a) Apparent clustering and apparent background earthquakes biased by undetected seismicity. *Journal of Geophysical Research* 110(B9): B09303.
Sornette D and Werner MJ (2005b) Constraints on the size of the smallest triggering earthquake from the epidemic-type aftershock sequence model, Båth's law, and observed aftershock sequences. *Journal of Geophysical Research* 110(B8): B08304.
Sornette D and Werner MJ (2009) Statistical physics approaches to seismicity. In: Meyers RA (ed.) *Encyclopedia of Complexity and Systems Science*, pp. 7872–7891. New York: Springer.
Stauffer D and Aharony A (1992) *Introduction to Percolation Theory*, 2nd edn. London: Taylor & Francis.
Sykes LR and Jaumé SC (1990) Seismic activity on neighboring faults as a long-term precursor to large earthquakes in the San Francisco Bay area. *Nature* 348(6302): 595–599.
Tiampo KF, Rundle JB, McGinnis S, Gross SJ, and Klein W (2002a) Eigenpatterns in southern California seismicity. *Journal of Geophysical Research* 107(B12): 2354.
Tiampo KF, Rundle JB, McGinnis S, Gross SJ, and Klein W (2002b) Mean-field threshold systems and phase dynamics: An application to earthquake fault systems. *Europhysics Letters* 60(3): 481–487.
Tiampo KF, Rundle JB, McGinnis SA, and Klein W (2002c) Pattern dynamics and forecast methods in seismically active regions. *Pure and Applied Geophysics* 159(10): 2429–2467.
Tiampo KF and Shcherbakov R (2012) Seismicity-based earthquake forecasting techniques: Ten years of progress. *Tectonophysics* 522: 89–121.
Tsapanos TM (1990) Spatial-distribution of the difference between the magnitudes of the main shock and the largest aftershock in the Circum-Pacific belt. *Bulletin of Seismological Society of America* 80(5): 1180–1189.
Turcotte DL (1997) *Fractals and Chaos in Geology and Geophysics*, 2nd edn. Cambridge: Cambridge University Press.
Turcotte DL (1999a) Seismicity and self-organized criticality. *Physics of the Earth and Planetary Interiors* 111(3–4): 275–293.
Turcotte DL (1999b) Self-organized criticality. *Reports on Progress in Physics* 62(10): 1377–1429.
Turcotte DL, Holliday JR, and Rundle JB (2007) BASS, an alternative to ETAS. *Geophysical Research Letters* 34(12), L12303.
Turcotte DL, Newman WI, and Shcherbakov R (2003) Micro and macroscopic models of rock fracture. *Geophysical Journal International* 152(3): 718–728.
Turcotte DL and Shcherbakov R (2006) Can damage mechanics explain temporal scaling laws in brittle fracture and seismicity? *Pure and Applied Geophysics* 163(5–6): 1031–1045.
Utsu T, Ogata Y, and Matsu'ura RS (1995) The centenary of the Omori formula for a decay law of aftershock activity. *Journal of Physics of the Earth* 43(1): 1–33.
Van Aalsburg J, Newman WI, Turcotte DL, and Rundle JB (2010) Record-breaking earthquakes. *Bulletin of Seismological Society of America* 100(4): 1800–1805.
Varnes DJ and Bufe CG (1996) The cyclic and fractal seismic series preceding an *m*(*b*) 4.8 earthquake on 1980 February 14 near the Virgin Islands. *Geophysical Journal International* 124(1): 149–158.
Vere-Jones D (1969) A note on the statistical interpretation of Båth's law. *Bulletin of Seismological Society of America* 59: 1535–1541.
Vere-Jones D (2009) Stochastic models for earthquake occurrence and mechanisms. In: Meyers RA (ed.) *Encyclopedia of Complexity and Systems Science*, pp. 2555–2580. New York: Springer.
Vere-Jones D (2010) Foundations of statistical seismology. *Pure and Applied Geophysics* 167(6–7): 645–653.
Vere-Jones D, Robinson R, and Yang WZ (2001) Remarks on the accelerated moment release model: Problems of model formulation, simulation and estimation. *Geophysical Journal International* 144(3): 517–531.
Ward SN (1992) An application of synthetic seismicity in earthquake statistics: The Middle America Trench. *Journal of Geophysical Research* 97(B5): 6675–6682.
Ward SN (1996) A synthetic seismicity model for southern California: Cycles, probabilities, and hazard. *Journal of Geophysical Research* 101(B10): 22393–22418.
Ward SN (2000) San Francisco Bay Area earthquake simulations: A step toward a standard physical earthquake model. *Bulletin of Seismological Society of America* 90(2): 370–386.
Wesnousky SG (1994) The Gutenberg–Richter or characteristic earthquake distribution, which is it? *Bulletin of Seismological Society of America* 84(6): 1940–1959.
Wissel F and Drossel B (2006) Transient and stationary behavior of the Olami–Feder–Christensen model. *Physical Review E* 74(6): 066109.
Yakovlev G, Gran JD, Turcotte DL, Rundle JB, Holliday JR, and Klein W (2010) A damage-mechanics model for fracture nucleation and propagation. *Theoretical and Applied Fracture Mechanics* 53(3): 180–184.
Yakovlev G, Newman WI, Turcotte DL, and Gabrielov A (2005) An inverse cascade model for self-organized complexity and natural hazards. *Geophysical Journal International* 163(2): 433–442.
Yakovlev G, Turcotte DL, Rundle JB, and Rundle PB (2006) Simulation-based distributions of earthquake recurrence times on the San Andreas Fault system. *Bulletin of Seismological Society of America* 96(6): 1995–2007.
Yamanaka Y and Shimazaki K (1990) Scaling relationship between the number of aftershocks and the size of the main shock. *Journal of Physics of the Earth* 38(4): 305–324.
Yang WZ, Vere-Jones D, and Li M (2001) A proposed method for locating the critical region of a future earthquake using the critical earthquake concept. *Journal of Geophysical Research* 106(B3): 4121–4128.
Yikilmaz MB, Heien EM, Turcotte DL, Rundle JB, and Kellogg LH (2011) A fault and seismicity based composite simulation in northern California. *Nonlinear Processes in Geophysics* 18(6): 955–966.
Yikilmaz MB, Turcotte DL, Yakovlev G, Rundle JB, and Kellogg LH (2010) Virtual California earthquake simulations: Simple models and their application to an observed sequence of earthquakes. *Geophysical Journal International* 180(2): 734–742.
Yoder MR, Holliday JR, Turcotte DL, and Rundle JB (2012a) A geometric frequency-magnitude scaling transition: Measuring $b = 1.5$ for large earthquakes. *Tectonophysics* 532: 167–174.
Yoder MR, Turcotte DL, and Rundle JB (2010) Record-breaking earthquake intervals in a global catalogue and an aftershock sequence. *Nonlinear Processes in Geophysics* 17(2): 169–176.
Yoder MR, Turcotte DL, and Rundle JB (2012b) Earthquakes: Complexity and extreme events. In: Sharma AS, Bunde A, Dimri VP, and Baker DN (eds.) *Extreme Events and Natural Hazards: The Complexity Perspective*. *Geophysical Monograph Series*, vol. 196, pp. 17–26. Washington, DC: AGU.
Yoder M, Van Aalsburg J, Turcotte D, Abaimov S, and Rundle J (2013) Statistical variability and Tokunaga branching of aftershock sequences utilizing BASS model simulations. *Pure and Applied Geophysics* 170: 155–171.
Zaliapin I, Gabrielov A, Keilis-Borok V, and Wong H (2008) Clustering analysis of seismicity and aftershock identification. *Physical Review Letters* 101(1): 018501.
Zöller G and Hainzl S (2001) Detecting premonitory seismicity patterns based on critical point dynamics. *Natural Hazards and Earth System Sciences* 1(1): 93–98.
Zöller G, Hainzl S, Ben-Zion Y, and Holschneider M (2009) Critical states of seismicity: From models to practical seismic hazard estimates space. In: Meyers RA (ed.) *Encyclopedia of Complexity and Systems Science*, pp. 7853–7872. New York: Springer.
Zöller G, Hainzl S, and Kurths J (2001) Observation of growing correlation length as an indicator for critical point behavior prior to large earthquakes. *Journal of Geophysical Research* 106(B2): 2167–2175.